

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2003, 09:19:25 ; Search time 1773.53 Seconds  
(without alignments)  
13149.734 Million cell updates/sec

Title: US-09-802-937-2

Perfect score: 1440

Sequence: 1 ctgcaggaagattaagg.....cgcgacgagtgctgac 1440

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: em\_estba.\*

2: em\_esthum.\*

3: em\_estin.\*

4: em\_estmu.\*

5: em\_estov.\*

6: em\_estpl.\*

7: em\_estro.\*

8: em\_htc.\*

9: gb\_est1.\*

10: gb\_est2.\*

11: gb\_htc.\*

12: gb\_est3.\*

13: gb\_est4.\*

14: gb\_est5.\*

15: em\_estfun.\*

16: em\_estom.\*

17: gb\_gss.\*

18: em\_gss\_hum.\*

19: em\_gss\_inv.\*

20: em\_gss\_pln.\*

21: em\_gss\_vrt.\*

22: em\_gss\_fun.\*

23: em\_gss\_mam.\*

24: em\_gss\_mus.\*

25: em\_gss\_other.\*

26: em\_gss\_pro.\*

27: em\_gss\_rod.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	99.4	6.9	637	17	AQ510538
2	99.4	6.9	813	17	AQ287155
3	98.4	6.8	627	17	AQ257106
4	95.8	6.7	488	17	AQ867664
5	95.8	6.7	718	17	AQ859314
6	95.2	6.6	757	17	AQ290644

7	94	6.5	1111	9	AB021959
c	8	93.6	6.5	749	17
	9	90.4	6.3	1101	17
	10	87.2	6.1	859	17
	11	86	6.0	746	17
	12	80.4	5.6	987	17
	13	79	5.5	1101	17
	14	77.8	5.4	1101	17
	15	77.8	5.4	1101	17
	16	77	5.3	1187	17
	17	76.6	5.3	1101	17
	18	76.4	5.3	634	17
	19	76.4	5.3	730	17
	20	75.2	5.2	1101	17
	21	75	5.2	734	17
	22	74.8	5.2	686	17
	23	72.8	5.1	543	17
	24	72.8	5.1	556	17
	25	70.6	4.9	776	17
	26	69.6	4.8	1101	17
	27	69	4.8	1101	17
	28	68.4	4.8	552	17
	29	68.4	4.8	1101	17
	30	68.2	4.7	868	17
	31	68.2	4.7	1101	17
	32	67.8	4.7	634	17
	33	67.4	4.7	1225	17
	34	67.2	4.7	994	17
	35	67.2	4.7	1101	17
	36	66.8	4.6	1101	17
	37	66.6	4.6	346	17
	38	66.2	4.6	1025	17
	39	66	4.6	1147	17
	40	65.8	4.6	1101	17
	41	65.6	4.6	609	17
	42	65.6	4.6	765	17
	43	65.4	4.5	1037	14
	44	65.4	4.5	1092	17
	45	65.4	4.5	1101	17

ALIGNMENTS

RESULT 1	AQ510538	AQ510538	637 bp	DNA	linear	GSS 04-MAY-1999
LOCUS	nbxb0095H17r	CUGI Rice BAC Library	Oryza sativa	genomic clone		
DEFINITION	nbxb0095H17r, DNA sequence.					
ACCESSION	AQ510538					
VERSION	AQ510538.1	GI:4733142				
KEYWORDS	GSS.					
SOURCE	Oryza sativa.					
ORGANISM	Oryza sativa					
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Euphorbia; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.					
AUTHORS	1 (bases 1 to 637)					
TITLE	Wing, R.A. and Dean, R.A.					
JOURNAL	A BAC End Sequencing Framework to Sequence the Rice Genome					
COMMENT	Unpublished (1998)					
	Contact: Wing RA					
	Clemson University Genomics Institute					
	Clemson University					
	100 Jordan Hall, Clemson, SC 29634, USA					
	Tel: 864 656 7288					
	Fax: 864 656 4293					
	Email: rwing@clemson.edu					
	Seq primer: GGAAACAGCTATGACCATG					
	Class: BAC ends					
	High quality sequence start: 95					
	High quality sequence stop: 424.					
	Location/Qualifiers					
FEATURES						

source

1. .637  
 /organism="Oryza sativa"  
 /strain="Japonica"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:4530"  
 /clones="nbx009SH17r"  
 /clone\_lib="CUGI Rice BAC Library"  
 /tissue\_type="Leaf"  
 /lab\_host="E. coli DH10B"  
 /note="Vector: pBelOBAC11; Site 1: HindIII; Site 2:  
 HindIII; Rice is one of two most popular grains in the  
 world. Half of the world population especially those  
 inhabiting highly populated areas of the humid tropics  
 and subtropics, rely on rice as their primary source of  
 carbohydrate. Monocotyledonous rice is a diploid plant  
 (2n=24) with a haploid genome equivalent of 431 Mbp  
 (Arumuganathan and Earle, 1991). The relatively small  
 genome of rice, three times larger than that of  
 Arabidopsis, makes it suitable for genomic studies. In  
 order to facilitate positional cloning, physical mapping  
 and genome sequencing of rice, we have constructed a BAC  
 library from Oryza sativa, Nipponbare variety. The  
 library contains 36,864 clones with an average insert size  
 of 128.5 Kb providing 10.9 haploid genome equivalents. The  
 deep coverage allows the isolation a particular sequence  
 with a probability of 99.9 %. Two high density filters,  
 each containing 18,432 clones (doubly spotted), represent  
 the whole library for colony screening."

BASE COUNT 207 a 134 c 121 g 173 t 2 others  
 ORIGIN  
 Query Match 6.9%; Score 99.4; DB 17; Length 637;  
 Best Local Similarity 74.6%; Pred. No. 6.8e-10;  
 Matches 138; Conservative 0; Mismatches 46; Indels 1; Gaps 1;  
 Qy 893 TTTCAAAACGCATGATAACGAGAAAGCTCTAGCACATTATTACTTAGATATTATTAAT 952  
 Db 423 TTTCCAGCACACGCAAAACGAGAAACTCTAGCACATGATTATTAAGTATTAACTAT 482  
 Qy 953 TATAAACTTGAAGAAAATATTAT-TTCAATTTTAAACAATGTATGCATAAATTATTT 1011  
 Db 483 TATAAAATTTGAAGAAAATAGATGTATTATTTATTAACAACCTCTATATAGAACTTT 542  
 Qy 1012 TTTAAACACACCAAAATTAACCCCTTTAAAAAGCATCTTAATAGGAAACGAGGAAGTTAA 1071  
 Db 543 TGTCAATAACGCACAAATTTAACGGTTTAGAAGCATGCTAACGGAAACGAGGAAGTTAA 602  
 Qy 1072 AGATT 1076  
 Db 603 AGTTT 607

RESULT 2  
 AQ287155 813 bp DNA linear GSS 03-DEC-1998  
 LOCUS nbx0026L24r CUGI Rice BAC Library Oryza sativa genomic clone  
 DEFINITION nbx0026L24r, DNA sequence.  
 ACCESSION AQ287155  
 VERSION AQ287155.1 GI:3948897  
 KEYWORDS GSS.  
 SOURCE Oryza sativa.  
 ORGANISM Oryza sativa

EuKaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzeae; Oryza.  
 1 (bases 1 to 813)  
 A BAC End Sequencing Framework to Sequence the Rice Genome  
 Unpublished (1998)  
 Contact: Wing RA  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA

REFERENCE  
 AUTHORS Wing,R.A. and Dean,R.A.  
 TITLE A BAC End Sequencing Framework to Sequence the Rice Genome  
 JOURNAL Unpublished (1998)  
 COMMENT Contact: Wing RA  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288  
 Fax: 864 656 4293  
 Email: rwing@clemson.edu  
 Seq primer: GGAACACAGCTATGACCATG  
 Class: BAC ends  
 High quality sequence stop: 473.  
 Location/Qualifiers

FEATURES  
 source

1. .813  
 /organism="Oryza sativa"  
 /strain="Japonica"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:4530"  
 /clones="nbx0026L24r"  
 /clone\_lib="CUGI Rice BAC Library"  
 /tissue\_type="Leaf"  
 /lab\_host="E. coli DH10B"  
 /note="Vector: pBelOBAC11; Site 1: HindIII; Site 2:  
 HindIII; Rice is one of two most popular grains in the  
 world. Half of the world population especially those  
 inhabiting highly populated areas of the humid tropics  
 and subtropics, rely on rice as their primary source of  
 carbohydrate. Monocotyledonous rice is a diploid plant  
 (2n=24) with a haploid genome equivalent of 431 Mbp  
 (Arumuganathan and Earle, 1991). The relatively small  
 genome of rice, three times larger than that of  
 Arabidopsis, makes it suitable for genomic studies. In  
 order to facilitate positional cloning, physical mapping  
 and genome sequencing of rice, we have constructed a BAC  
 library from Oryza sativa, Nipponbare variety. The  
 library contains 36,864 clones with an average insert size  
 of 128.5 Kb providing 10.9 haploid genome equivalents. The  
 deep coverage allows the isolation a particular sequence  
 with a probability of 99.9 %. Two high density filters,  
 each containing 18,432 clones (doubly spotted), represent  
 the whole library for colony screening."

BASE COUNT 259 a 182 c 138 g 234 t  
 ORIGIN  
 Query Match 6.9%; Score 99.4; DB 17; Length 813;  
 Best Local Similarity 74.6%; Pred. No. 6.5e-10;  
 Matches 138; Conservative 0; Mismatches 46; Indels 1; Gaps 1;  
 Qy 893 TTTCAAAACGCATGATAACGAGAAAGCTCTAGCACATTATTACTTAGATATTATTAAT 952  
 Db 369 TTTCCAGCACACGCAAAACGAGAAACTCTAGCACATGATTATTAAGTATTAACTAT 428  
 Qy 953 TATAAACTTGAAGAAAATA-TTTATTTGAATTTTAAACAATGTATGCATAAATTATTT 1011  
 Db 429 TATAAAATTTGAAGAAAATAGATTTATTTATTTATTAACAACCTCTATATAGAACTTT 488  
 Qy 1012 TTTAAACACACCAAAATTAACCCCTTTAAAAAGCATCTTAATAGGAAACGAGGAAGTTAA 1071  
 Db 489 TGTCAATAACGCACAAATTTAACGGTTTAGAAGCATGCTAACGGAAACGAGGAAGTTAA 548  
 Qy 1072 AGATT 1076  
 Db 549 AGTTT 553

RESULT 3

AQ257106

LOCUS

DEFINITION

nbx0017B06r CUGI Rice BAC Library Oryza sativa genomic clone

nbx0017B06r, DNA sequence.

ACCESSION

AQ257106

VERSION

AQ257106.1 GI:3781588

KEYWORDS

GSS.

SOURCE

ORGANISM

Oryza sativa.

EuKaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzeae; Oryza.  
 1 (bases 1 to 627)

AUTHORS Wing, R.A. and Dean, R.A.  
TITLE A BAC End Sequencing Framework to Sequence the Rice Genome  
JOURNAL Unpublished (1998)  
COMMENT Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
Seq primer: GGAAACAGCTATGACCATG  
Class: BAC ends  
High quality sequence start: 4  
High quality sequence stop: 368.  
Location/Qualifiers  
1. 627  
/organism="Oryza sativa"  
/strain="Japonica"  
/cultivar="Nipponbare"  
/db\_xref="taxon:4530"  
/clone="nbx0017B06r"  
/clone\_lib="CUGI Rice BAC Library"  
/tissue\_type="leaf"  
/lab\_host="E. coli DH10B"  
/notes="Vector: pBelBAC11; Site 1: HindIII; Site 2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."

BASE COUNT 202 a 135 c 116 g 173 t 1 others  
ORIGIN

Query Match 6.8%; Score 98.4; DB 17; Length 627;  
Best Local Similarity 74.1%; Pred. No. 1.1e-09;  
Matches 137; Conservative 0; Mismatches 47; Indels 1; Gaps 1;

Qy 893 TTTCAAACGCATGATAAACGAGAAAGCTCATTAGCACATTATTACTTTAGATATTTATAAT 952  
Db 387 TTTTCAGCACACGCAAAACGAGAAAGCTCATTAGCACATGATTAATTAAGTATTACTAT 446

Qy 953 TATAAACTTGAAAAAATA-TTTATTGAAATTTTAAAACAATGTATGCAATAATTTT 1011  
Db 447 TATAAAAATTGAAAAAATAGATTTATTTAATTTTATAAACAACCTCTATATAAAAACTNT 506

Qy 1012 TTTTAAACACACCAATTTAACCCCTTTAAAAGCATCCTTAATAGGAACGAGGAGTTAA 1071  
Db 507 TGTCAATACGCACAATTTAACCGTTTAGAAGCATCTCTAAACGAAAAACGAGGAGTTTA 566

Qy 1072 AGATT 1076  
Db 567 AGTTT 571

RESULT 4  
AQ867664 488 bp DNA linear GSS 03-NOV-1999  
LOCUS nbx0033C04f CUGI Rice BAC Library (EcoRI) Oryza sativa genomic  
DEFINITION clone nbx0033C04f, DNA sequence.  
ACCESSION AQ867664

VERSION AQ867664.1 GI:6218121  
KEYWORDS GSS.  
SOURCE Oryza sativa.  
ORGANISM Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.  
REFERENCE 1 (bases 1 to 488)  
AUTHORS Wing, R.A. and Dean, R.A.  
TITLE A BAC End Sequencing Framework to Sequence the Rice Genome  
JOURNAL Unpublished (1998)  
COMMENT Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
Seq primer: TAATACGACTCACTATAGGG  
Class: BAC ends  
High quality sequence start: 9  
High quality sequence stop: 461.  
Location/Qualifiers  
1. 488  
/organism="Oryza sativa"  
/strain="Japonica"  
/cultivar="Nipponbare"  
/db\_xref="taxon:4530"  
/clone="nbx0033C04f"  
/clone\_lib="CUGI Rice BAC Library (EcoRI)"  
/tissue\_type="leaf"  
/lab\_host="E. coli DH10B"  
/notes="Vector: pBACIndigo; Site 1: EcoRI; Site 2: EcoRI; Rice is the most important food crop in the world. Half of the world population, especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety using EcoRI as the cloning enzyme. The library contains 55,296 clones with an average insert size of 121 Kb providing approximately 15 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Three high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening and can be requested from the Clemson University BAC/EST Resource Center (www.genome.clemson.edu)."

BASE COUNT 177 a 74 c 83 g 154 t  
ORIGIN

Query Match 6.7%; Score 95.8; DB 17; Length 488;  
Best Local Similarity 73.4%; Pred. No. 3.7e-09;  
Matches 149; Conservative 0; Mismatches 52; Indels 2; Gaps 2;

Qy 874 CCTCTAGTTGGAGGTGATTTTCAAACGCATGATAAACGAGAAAGCTCATTAGCACATTA 933  
Db 104 CATCCAGTTGGAGGTATATTTTC-AGCACATGCAAAACGAATGAACCTATTAGCATGA 162

Qy 934 TTACTTAGATATTTATTAATAACTTGAAAAAATATTTATTGAAATTTTTTAAACAA 993  
Db 163 TTAATTAATATTAATTAATAAATGAAAAATAGTATTATTATTTTAACTACT 222

Qy 994 TGTATGCATAAATTTATTTTAAAAACACACCAATTTTAAACCCCTTTAAAAAGCATCCTTAAT 1053  
Db 223 TATACATATAAACITTTTATCAAAAATACACC-ATTTAATAGTTTAAATAACATGCTAAC 281

Qy 1054 AGGAAACGAGGAAGTTAAAGATT 1076

Dn	282	GGAACAGGAAATTTTCAAATT	304		
RESULT 5	AQ859314	718 bp	DNA	linear GSS 03-NOV-1999	
LOCUS	nbeb0012C12r	CUGI Rice BAC Library (EcoRI)	Oryza sativa genomic clone nbeb0012C12r,	DNA sequence.	
DEFINITION	AQ859314	GI:6209771			
ACCESSION	AQ859314.1	GI:6209771			
VERSION	GSS:				
KEYWORDS	Oryza sativa.				
SOURCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoidae; Oryzeae; Oryza.				
ORGANISM	1 (bases 1 to 718)				
REFERENCE	Wing,R.A. and Dean,R.A. A BAC End Sequencing Framework to Sequence the Rice Genome Unpublished (1998) Contact: Wing RA Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 7288 Fax: 864 656 4293 Email: rwing@clemson.edu Seq primer: GGAAACGGTAGGCACATG Class: BAC ends High quality sequence start: 15 High quality sequence stop: 440.				
FEATURES	Location/Qualifiers				
source	1..718	/organism="Oryza sativa"			
		/strain="Japonica"			
		/cultivar="Nipponbare"			
		/db_xref="taxon:4530"			
		/clone_lib="nbeb0012C12r"			
		/tissue_type="Leaf"			
		/lab_host="E. coli DH10B"			
		/note="Vector: pBlacIndigo; Site 1: EcoRI; Site 2: EcoRI; Rice is the most important food crop in the world. Half of the world population, especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety using EcoRI as the cloning enzyme. The library contains 55,296 clones with an average insert size of 121 kb providing approximately 15 haploid genomes equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Three high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening and can be requested from the Clemson University BAC/EST Resource Center ( <a href="#">www.genome.clemson.edu</a> )."			
BASE COUNT	247 a	127 c	131 g	211 t 2 others	
ORIGIN					
Query Match	6.7 %;	Score 95.8;	DB 17;	Length 718;	
Best Local Similarity	73.4 %;	Pred. No. 3.4e-09;			
Matches 149;	Conservative	0;	Mismatches 52;	Indels 2;	Gaps 2;
Qy	874	CCCTCACTTGCGAGGTTCATTTTTTAACAACCAGCATATAAACAGAAGACTCATTAGCACATTA	933		
b	151	CATCCAGTTGGAGGTATATTTTC-AGCATATCAAAAACGAATGAACCTATTAGCACATGA	209		



```

Query Match      6.6%; Score 95.2; DB 17; Length 757;
Best Local Similarity 73.0%; Pred. No. 4.5e-09;
Matches 135; Conservative 0; Mismatches 49; Indels 1; Gaps 1;

QY 893 TTTCACAGCATGATAACGAGAAAGCTATTAGCACATTTACTAGTATTTTATAT 952
Db 370 TTTCAGCACACGAAACGAGAAAGCTACTAGCACATGATTAAATTAAGTATTAACAT 429

QY 953 TATAAACTTGAATAAATA-TTTATTGTAATTTTTTAAACAATGTATGCATAAATTTAT 1011
Db 430 TATAAAATGGAATAAATAAGATTATTTATTTATTATAGACAATCTTCTATATAGAACTTT 489

QY 1012 TTTAAACAACACAAATTTAACCTTTTAAAGATCCTTAATAGGAAACGAGGAAGTTAA 1071
Db 490 TGTCAATAAGCCACAATNAACGGTTTAGAAGCATGCTAACGGAACGAGGAAGTTAA 549

QY 1072 AGATT 1076
Db 550 AGTTT 554

RESULT 7
AB021959
LOCUS      AB021959      1111 bp      mRNA      linear      EST 07-JAN-1999
DEFINITION      Oryza sativa ms-bo-Taichung 65 immature anther Oryza
ACCESSION      AB021959
VERSION      AB021959.1 GI:4107146
KEYWORDS      EST.
SOURCE      Oryza sativa.
ORGANISM      Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE      1 (bases 1 to 1111)
AUTHORS      Toriyama,K. and Ohnashi,Y.
TITLE      EST of immature anther of rice
JOURNAL      Unpublished (1999)
COMMENT      Contact: Toriyama K
              Faculty of Agriculture, Laboratory of Plant Breeding
              Tohoku University
              Tsutsumidori-Anamiyamachi 1-1, Aobaku, Sendai, Miyagi 981-8555,
              Japan
              Email: torikin@bios.tohoku.ac.jp.
              Location/Qualifiers
                1..1111
                  /organism="Oryza sativa"
                  /strain="ms-bo-Taichung 65"
                  /db_xref="taxon:4530"
                  /clone="BD9R6-1-7"
                  /clone_lib="Oryza sativa ms-bo-Taichung 65 immature
                  anther"
                  /tissue_type="anther"
                  /dev_stage="immature anther"
BASE COUNT      335 a 184 c 265 g 327 t
ORIGIN

Query Match      6.5%; Score 94; DB 9; Length 1111;
Best Local Similarity 65.9%; Pred. No. 7.2e-09;
Matches 153; Conservative 0; Mismatches 75; Indels 4; Gaps 1;

QY 849 AAACACAGCCTTAAGGCTTCTGTAGTCTCTAGTTGGAGGTTGATTTTCAACGCGATGATA 908
Db 804 AAACATCCAGGCTGCGTTTCACTACACTGTGCAGGTAGATTTTTTTAGTGTCATATA 863

QY 909 AACGAGAAAGCTCATTAGCACATTTACTTAGATATTTTATAATTTAAACCTTGAAAAA 968
Db 864 AACGAGAAAGCTTATTAGCACATGATTAAATTAAGTATTAATAATTTATTTTAAAC 923

QY 969 ATATTTATTTGAATTTTTTAAACATGTCATGATAAATTTATTTTAAACACACCA-- 1026
Db 924 AGATTTGTTGAATTTTTTAAACAACTTCTATATAGACTTTTTTTTTTAAAAAAAATACA 983

```

---

```

QY 1027 --ATTTAACCTTTTAAAGCATCCTTAATAGGAACGAGGAAGTTAAAGATT 1076
Db 984 TCATGTACAGTTTAAAAAATCGTGTATATAAGAAACCGAGAAGTTGAAGTTT 1035

RESULT 8
AQ160799/c
LOCUS      AQ160799      749 bp      DNA      linear      GSS 12-SEP-1998
DEFINITION      nbxb0006C07f CUGI Rice BAC Library Oryza sativa genomic clone
ACCESSION      nbxb0006C07f, DNA sequence.
VERSION      AQ160799
KEYWORDS      AQ160799.1 GI:3552496
SOURCE      GSS.
ORGANISM      Oryza sativa.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE      1 (bases 1 to 749)
AUTHORS      Wing,R.A. and Dean,R.A.
TITLE      A BAC End Sequencing Framework to Sequence the Rice Genome
JOURNAL      Unpublished (1998)
COMMENT      Contact: Wing RA
              Clemson University Genomics Institute
              Clemson University
              100 Jordan Hall, Clemson, SC 29634, USA
              Tel: 864 656 7288
              Fax: 864 656 4293
              Email: rwing@clemson.edu
              Seq primer: TAATGACACTCACTATAGG
              Class: BAC ends
              High quality sequence stop: 464.
              Location/Qualifiers
                1..749
                  /organism="Oryza sativa"
                  /strain="Japonica"
                  /cultivar="Nipponbare"
                  /db_xref="taxon:4530"
                  /clone="nbxb0006C07f"
                  /clone_lib="CUGI Rice BAC Library"
                  /tissue_type="Leaf"
                  /lab_host="E. coli DH10B"
                  /note="vector: pBelOBAC11; Site 1: HindIII; Site 2:
                  HindIII; Rice is one of two most popular grains in the
                  world. Half of the world population especially those
                  inhabiting highly populated areas of the humid tropics
                  and subtropics, rely on rice as their primary source of
                  carbohydrate. Monocotyledonous rice is a diploid plant
                  (2n=24) with a haploid genome equivalent of 431 Mbp
                  (Arumuganathan and Earle, 1991). The relatively small
                  genome of rice, three times larger than that of
                  Arabidopsis, makes it suitable for genomic studies. In
                  order to facilitate positional cloning, physical mapping
                  and genome sequencing of rice, we have constructed a BAC
                  library from Oryza sativa, Nipponbare variety. The
                  library contains 36,864 clones with an average insert size
                  of 128.5 Kb providing 10.9 haploid genome equivalents. The
                  deep coverage allows the isolation a particular sequence
                  with a probability of 99.9 %. Two high density filters,
                  each containing 18,432 clones (doubly spotted), represent
                  the whole library for colony screening."
BASE COUNT      252 a 145 c 114 g 232 t 6 others
ORIGIN

Query Match      6.5%; Score 93.6; DB 17; Length 749;
Best Local Similarity 68.8%; Pred. No. 9.4e-09;
Matches 139; Conservative 0; Mismatches 60; Indels 3; Gaps 1;

QY 878 TAGTTGGAGGTTGATTTTCAACGCGATGATAACGAGAAAGCTCATGACATTTATAC 937
Db 744 TAGTTGNAGNAGCTTTTTCAGNCCGCTGTGAATGAGANAGCCCATGACATGATTTAA 685

QY 938 TTAGATATTTATTAATTATAAACTTGAAAAAATAATTTTATTGAAATTTTTT---AAACAAT 994

```

```

684 TTAAGTATTAACTATTATANACTTGAATAATGGATTATTATTGATTTTAAAAAAGC 625
QY 995 GTATGCAATTAATTTTAAACACACCAAAATTAACCCCTTAAAGACATCCTTAATA 1054
Db 624 TTTTCTATAAAAGTTTTTGCAGAAACCGCCCACTTAAATAGTNTGAAAGTGTCTGATG 565
QY 1055 GGAACGAGGAAGTTAAAGATT 1076
Db 564 GAAACCAAGGAGTTGATGTTT 543

RESULT 9
CNS00EVL 1101 bp DNA linear GSS 04-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC:
DEFINITION BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL069706
VERSION AL069706.1 GI:4949849
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
REFERENCE
AUTHORS
TITLE Direct Submission
JOURNAL
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.
FEATURES
source
1..1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR29B23"
/clone_lib="RPCI-98"
/notes="end : T7"
BASE COUNT 419 a 91 c 60 g 299 t 232 others
ORIGIN
Query Match 6.3%; Score 90.4; DB 17; Length 1101;
Best Local Similarity 36.2%; Pred. No. 3.7e-08;
Matches 228; Conservative 107; Mismatches 285; Indels 10; Gaps 2;

574 ACAACCACTGTTTCGTACTCACTCTAATTTGTAAATTCCTATTTCAGTCACAAAATTC 633
Db 445 MMMHMAATYCTCAHTTMMMMHMAATTTWAAAWAAATTTATWATWAAAWAAW 504
QY 634 CAATTTCAATTAAGAAAAATAAACGTAGACGGCTAACCCCAATCAAGGCTAAGTT 693
Db 505 WWAATTTTMMWWTWTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 564
QY 694 CGAGAGGTGAGTACGCAGCAAAATATGATGGTTTATATATCA-----TTTTTT 745
Db 565 TTAWAAATTTTAAAWWTTATATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 624

```

```

QY 746 TTAATAACTTTTCAATAAAATTTCTTTAGGAAACATATCATTTTAATGGTTGAAAAACGT 805
Db 625 ATAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 684
QY 806 GCACATTAAGAAACTAAGACGATGAGTTGGGAAACAAGAGAAAAACACAGCCTTAAGGC 865
Db 685 AAATAWAAAAAATAAAAAAATAAAATWAAATWATAAAATWATAAAATWATAAAATWATA 744
QY 866 TTCTTGATCTCTAGTTGGAGGTTGATTTTCAACGCGATGATAAATCAAGAAAGCTCATTA 925
Db 745 ATWAWATAATATWATAATATATATTTTAAWWTATWAAWWTATATWATAWATAWAAWAW 804
QY 926 GCACATTATTACTTATGATATTTATAATATAAATCTGAAAAAAATATTTATTTGAAATTT 985
Db 805 ATAAATATWATAWATAWAAWAAWATAWATAWATAWATAWATAWATAWATAWATAWATAW 864
QY 986 TTAACCAATGATGCATAAATTA--TTTTTTAAAAACACACCAATTTTAAACCTTTAAAAA 1043
Db 865 AATWATAWAAAAAATAAAATWAAWTTTWTWTWTWTWAAWATAAATAWATAWAAAAAATA 924
QY 1044 GCATCCTAATAGGAAACGAGGAGCTTAAAGATTCCACCGAAGTGTTCGATATGAAAAAT 1103
Db 925 AAAAAATPAWAAWWTATATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 984
QY 1104 GGGGTGGATTGAAATTTGGTAAATGAAATCAGGTTAGGATTAAATATATAAATGAAAGAG 1163
Db 985 ATWTATATATWTTATWATAWATAATTTTAAWWTATATTTTAAWAAWATAATATATATWATA 1044
QY 1164 GGAGAAATGAATGGTTAGGTTTAAATGCT 1193
Db 1045 WWTAWATATAWAAWAAATTTAAATTTATATATAT 1074

RESULT 10
AZ045395/c 859 bp DNA linear GSS 08-MAR-2000
LOCUS nbe0081C21f CUGI Rice BAC Library (EcoRI) Oryza sativa genomic
DEFINITION clone OSJNB0081C21f, DNA sequence.
ACCESSION AZ045395
VERSION AZ045395.1 GI:7207135
KEYWORDS GSS.
SOURCE Oryza sativa.
ORGANISM Oryza sativa.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 859)
REFERENCE
AUTHORS Wing,R.A. and Dean,R.A.
TITLE A BAC End Sequencing Framework to Sequence the Rice Genome
JOURNAL Unpublished (1998)
COMMENT Contact: Wing RA
Clemson University Genomics Institute
Clemson University,
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: GTAAACGACGCGCAGTG
Class: BAC ends
High quality sequence start: 132
High quality sequence stop: 324.
FEATURES
source
1..859
/organism="Oryza sativa"
/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/clone="OSJNB0081C21f"
/clone_lib="CUGI Rice BAC Library (EcoRI)"
/lab_host="E. coli DH10B"
/notes="vector: pBACIndigo; Site_1: EcoRI; Site_2: EcoRI;
Rice is the most important food crop in the world. Half of

```

the world population, especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety using EcoRI as the cloning enzyme. The library contains 55,296 clones with an average insert size of 121 Kb providing approximately 15 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Three high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening and can be requested from the Clemson University BAC/EST Resource Center (www.genome.clemson.edu).

BASE COUNT 217 a 231 c 160 g 247 t 4 others  
ORIGIN

Query Match 6.1%; Score 87.2; DB 17; Length 859;  
Best Local Similarity 62.6%; Pred. No. 1.7e-07;  
Matches 169; Conservative 0; Mismatches 98; Indels 3; Gaps 2;

QY 812 AAGAAACTAAGAACGATGAGTTGGAAACAAGAGAAACACAGCCTTAAGGCTTCTTG 871  
DB 574 AAAAAAGAAAGGGTCACACATGACACACCACTAAACACACCTCTAAGGCCCGTTC 515  
QY 872 ATCCTCTAGTTGGAGGTGA-TTTTCAACGCATGATAAACGAGAAAGCTCATTAGCACA 930  
DB 514 GTTCTCTCCGATTAGAGGGACATTTCCGGTGCACATGAAATGAGAAACCTCATTAGCACA 455  
QY 931 TTATTACTTAGATATTATTAATTAATACTTGAAAAAATATTTATTGAAATTTTAA 990  
DB 454 TAATTAATTAAGTATTAAATATTATTAATAATTAATAATTAATAATTTATTCGATTTTAA 395  
QY 991 CAA-TGTATGCAATAATTATTTTAAAAACACACCAATTTTAACCTTTTAAAAAGCATC 1048  
DB 394 TAAGTTTACGTAGAACTTTTGGCCAAAACACACATTTACCAGTTTGAANAACGTGC 335  
QY 1049 CTAATAGAAACGAGGAAGTTAAAGATTCA 1078  
DB 334 TTACGGAACACGAGGAGTCAAGTTGA 305

RESULT 11  
LOCUS AQ160781/c 746 bp DNA linear GSS 12-SEP-1998  
DEFINITION nbx0006A07E CUGI Rice BAC Library Oryza sativa genomic clone  
ACCESSION nbx0006A07E, DNA sequence.  
VERSION AQ160781  
KEYWORDS GSS.  
SOURCE Oryza sativa.  
ORGANISM Oryza sativa.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.  
1 (bases 1 to 746)

REFERENCE  
AUTHORS Wing, R.A. and Dean, R.A.  
TITLE A BAC End Sequencing Framework to Sequence the Rice Genome  
JOURNAL Unpublished (1998)  
COMMENT Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
Seq primer: TAATACCACTCACTATAGGG  
Class: BAC ends

High quality sequence stop: 450.  
Location/Qualifiers  
1. .746

/organism="Oryza sativa"  
/strain="Japonica"  
/cultivar="Nipponbare"  
/db\_xref="taxon:4530"  
/clone="nbx0006A07E"  
/clone\_lib="CUGI Rice BAC Library"  
/tissue\_type="Leaf"  
/lab\_host="E. coli DH10B"  
/note="Vector: pBeloBAC11; Site 1: HindIII; Site 2:  
HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 55,296 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."

BASE COUNT 249 a 146 c 115 g 235 t 1 others  
ORIGIN

Query Match 6.0%; Score 86; DB 17; Length 746;  
Best Local Similarity 70.6%; Pred. No. 3e-07;  
Matches 142; Conservative 0; Mismatches 56; Indels 3; Gaps 2;

QY 878 TAGTTGGAGGTGATTTTCAAACGCATGATAAACGAGAAAGCTCATTAGCACATTATTAC 937  
DB 742 TAGTTGAAGAAGCTTTTCAGCGC-TCGGAATGAGAAAGCCCATTAGCACATGATTAA 684  
QY 938 TTAGATATTATTAATAATAAACTTGAAAAAATATTTATTG--AATTTTAAACAATG 995  
DB 683 TTAGCTATTAACTATTATAAACTTGAAAAATGGATNTATTTGATATTTTAAAAAACA 624  
QY 996 TATGCATAAATTATTTTAAAAACACACCAATTTAACCTTTTAAAAAGCATCTAATAG 1055  
DB 623 TTTTCTATAAAAGTTTTCGAAAACGGCGCCATTTAATAGTTTGAAAAGTGTCTGATGG 564  
QY 1056 GAAACGAGGAAGTTAAAGATT 1076  
DB 563 AAACAAGGAAGTTGTAGTTT 543

RESULT 12  
LOCUS CDS014PQ 987 bp DNA linear GSS 26-JUL-1999  
DEFINITION Drosophila melanogaster genome survey sequence SP6 end of BAC  
BACN12P22 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.  
ACCESSION AL104456  
VERSION AL104456.1 GI:5616067  
KEYWORDS GSS.  
SOURCE Drosophila melanogaster.  
ORGANISM Drosophila melanogaster.

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 987)

REFERENCE  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr



```

QY 836 GGAACAAGAGAAACACAGCCTTAAGGCTTCITGATCCTCTAGTGGAGGTTGATTT 895
Db 562 TATTTTAAARAAGAAWAAWATTTATTTTWTWTWTWTWTWTWTWTWTWTWTAAWAAWTA 503
QY 896 CAAACGCATGATAAAGCAGAAAGCTCATTAGCACATTTACTTAGATATTATAATAT 955
Db 502 TWAA---WTAAGAAAAAATAATTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 446
QY 956 AAACCTGAAAAAATATTATTGTAATTTTAAACAATGTAT 998
Db 445 AATTTTWTWTWAAATTTATTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 403

RESULT 14
CNS00E07
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC;
BACR29P01 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL069440.1 GI:4949583
VERSION
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
Bp 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osogawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
source
1. .1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR29P01"
/clone_lib="RPCI-98"
/note="end : TET3"
BASE COUNT 366 a 66 c 104 g 351 t 214 others
ORIGIN
Query Match 5.5%; Score 79; DB 17; Length 1101;
Best Local Similarity 37.3%; Pred. No. 6.9e-06;
Matches 176; Conservative 67; Mismatches 229; Indels 0; Gaps 0;

QY 718 ATATGATGTTTATTAATGATTTTAAATAACTTTCACATAATTTCTTTAGGAA 777
Db 534 AAATAATTTTAAATAAATTTTATTTTATTTTAAATAAATAATTTTAAATTTTAA 593
QY 778 ACATATCATTTAATGTTTGAAGAAACGTCACATAAATAACATAAGCATGCTGG 837
Db 594 TTTTATATATTAAAGWAAAGAAAAAATTAATTAATTTTAAATTTTAAATTTTAA 653
QY 838 AAACAAGAGAAAAACAGCCTTAAAGCTTCTTGATCCTCTAGTGGAGGTTGATTTCA 897
Db 654 TAAATTTTAAATTTAAATTTTAAWATTTTAAWATTTTAAATTTTAAARWAAATTTAA 713

```

```

QY 898 AACGCATGATTAACAGAGAGCTCATTAGCACATTTACTTAGATATTATAATATA 957
Db 714 TTTWAATAAATAAATAAATAAATTTTATTTTAAWAAWTTAAATTTATTTAAWTA 773
QY 958 ACTTGAAAAAATAATTTTATTTGAAATTTTAAACAATGTATGCATAAATTTTAA 1017
Db 774 WTTTAAWAAAAAATAATTTTAAWATTTTAAWATTTTAAWATTTTAAWATTTAA 833
QY 1018 AACACACCAATTTAACCCCTTTAAAGAGCATCTTAATAGGAACGAGGAGTTAA 1077
Db 834 TTTWAATATTTTATTTTAAWATTTTAAWATTTTAAWATTTTAAWATTTAAW 893
QY 1078 ACCGAAGTCTTTGGATATGAATAATGGGTGGATTTAGAAATTTGAATGAATCAG 1137
Db 894 AAAAATAATGTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 953
QY 1138 TAGGATTAATTTAAATTTAAAGAGGAGGAATGAATGGTTAGAGTTTAAAT 1189
Db 954 TATGTATAWAAATTTWTGTATTAAATAATATGATATTTAAATAAAAAAAT 1005

RESULT 15
CNS001FB/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR04A23 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL060732.1 GI:4939397
VERSION
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
Bp 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osogawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
source
1. .1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR04A23"
/clone_lib="RPCI-98"
/note="end : TET3"
BASE COUNT 288 a 110 c 103 g 491 t 109 others
ORIGIN
Query Match 5.4%; Score 77.8; DB 17; Length 1101;
Best Local Similarity 38.3%; Pred. No. 1.2e-05;
Matches 215; Conservative 62; Mismatches 284; Indels 1; Gaps 1;

QY 433 TAGCATTTTCTTTTCAAAAAGAAATTTATTAATTTTCTTATGAACGCAATTAACC 492

```

Db 1083 TATAATWTATATWAAAAATWAATAAAATATWAAATWATAWATAATAAAAAATATAAAATW 1024  
Qy 493 GTTCGAGAAATATGCTGTCATAAATAGTAGTCTAGTCGAGAAACAAATTTAATATCACA 552  
Db 1023 ATTAAAAATAAATATWAAAAATWAAAAATWAAAAATTTWAATTAATTATATATTTNAAAAA 964  
Qy 553 TAAAAAAGAGGTTGTAAATTACAAACCATGTTTCGTACTACAACTCTAAATTTGTAAAT 612  
Db 963 AAAAAAATAATTAATAWAAAAAATAAATAATAATAATAATAATAATAATAATAATAATA 904  
Qy 613 CTTATTTAGTCACAAAATTCCAAATTCCAAATTAAGAAA-AAATAACGTAGACGGCTAAG 671  
Db 903 ATTAAAAAATAATAWAAATTAWAAAAATAWAAAAATAWAAAAATAWAAAAATAWAAAT 844  
Qy 672 CCCACCATCTAAGGCTAAGTTTCGAGAGGTGAAGTACGCACGAAATAATGATGTTTAT 731  
Db 843 WAAAAATWNTWAAAAATWTTTTTTAATAATAATAATAATAATAATAATAATAATAATA 784  
Qy 732 TAATATGATTTTTTTTAAATACTTTACATAAAATTTCTTTAGGAAACATATCATTTAAT 791  
Db 783 ATAAAAAATWNTWAAAAATAAATAATAATAATAATAATAATAATAATAATAATAATA 724  
Qy 792 GGTTCAAAAACGTGCACATAGAAAACCTAAGAACGATCGAGTTGGGAAACAAGAGAAAA 851  
Db 723 WWTAAAAATAAAAAATAWATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 664  
Qy 852 CACAGCCTTAAGGCTTCTTGATCCTCTAGTTGGAGGTTGATTTTCAACGCATGATAAAC 911  
Db 663 AAAAGNNGGGGGGGGGGGGGGGGGGGHYYWTHWHTATNTTATATATTA 604  
Qy 912 GAGAAAGCTCATTAGCACATTATTACTTAGATATTATTAATTAATACTTGAAAAAATA 971  
Db 603 TTATTACATTATTATTATNTTACTATTATTATTATTATTATTATTATTATTATTATT 544  
Qy 972 TTTATTGAATTTTTTAAACAA 993  
Db 543 TTAATACTATATTTCAACCCA 522

Search completed: March 15, 2003, 16:52:51  
Job time : 1786.53 secs

GenCore version 5.1.4.p5 4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2003, 09:19:25 ; Search time 5346.47 Seconds  
(without alignments)  
13149.734 Million cell updates/sec

Title: US-09-802-937-1

Perfect score: 4341

Sequence: 1 ctgcaggaagattaattagg.....gtatatcttcattgttttt 4341

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: em\_estba.\*
- 2: em\_esthum.\*
- 3: em\_estin.\*
- 4: em\_estmu.\*
- 5: em\_estov.\*
- 6: em\_estpl.\*
- 7: em\_estro.\*
- 8: em\_htc.\*
- 9: gb\_est1.\*
- 10: gb\_est2.\*
- 11: gb\_htc.\*
- 12: gb\_est3.\*
- 13: gb\_est4.\*
- 14: gb\_est5.\*
- 15: em\_estfun.\*
- 16: em\_estom.\*
- 17: gb\_gss.\*
- 18: em\_gss\_hum.\*
- 19: em\_gss\_inv.\*
- 20: em\_gss\_pin.\*
- 21: em\_gss\_vrt.\*
- 22: em\_gss\_fun.\*
- 23: em\_gss\_mam.\*
- 24: em\_gss\_mus.\*
- 25: em\_gss\_other.\*
- 26: em\_gss\_pro.\*
- 27: em\_gss\_rod.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	448	10.3	680	17	AQ257253 nbxb0017P
2	439.6	10.1	520	17	AQ289489 nbxb0035C
3	238	5.5	502	17	AQ915399 nbebo056H
4	173.2	4.0	287	9	AU164649 AU164649
5	156.4	3.6	822	17	BH704748 BOMOR39TR
6	149.6	3.4	648	12	BF473637 WHE0930_F

7	149.6	3.4	680	10	BE493591
8	148.6	3.4	514	10	BE405815
9	148	3.4	774	14	B0840903
10	147.6	3.4	596	10	BE443438
11	146.4	3.4	612	14	BQ243888
12	146.4	3.4	719	14	BQ246627
13	145.2	3.3	864	12	BG418644
14	144.8	3.3	593	10	BE495096
15	144.8	3.3	639	10	BE403629
16	143.2	3.3	420	9	AJ486371
17	143.2	3.3	596	14	B0753292
18	143.2	3.3	599	13	BM098903
19	143.2	3.3	600	10	AV929110
20	143.2	3.3	603	10	AV913445
21	143.2	3.3	634	10	AV924023
22	143.2	3.3	644	10	AV946226
23	143.2	3.3	648	14	BQ458847
24	143.2	3.3	655	13	BJ463888
25	143.2	3.3	679	10	AV918748
26	143.2	3.3	742	13	BJ468059
27	143.2	3.3	746	10	BE421072
28	143.2	3.3	906	12	BG301342
29	142.2	3.3	560	13	BJ467426
30	141.6	3.3	518	13	BJ459088
31	141.6	3.3	531	13	BJ451593
32	141.6	3.3	811	12	BF621279
33	138	3.2	404	14	BM953435
34	137.2	3.2	482	14	BQ236703
35	136.4	3.1	564	14	BQ163113
36	136.4	3.1	585	14	BQ528230
37	136.4	3.1	582	9	AI670171
38	136.4	3.1	583	12	BF587853
39	136.4	3.1	602	14	BM95705
40	136.4	3.1	613	14	BQ294265
41	136.4	3.1	657	9	AI739940
42	136.4	3.1	1088	11	AY109686
43	136.2	3.1	445	14	BQ239883
44	135.2	3.1	508	17	BH693012
45	134	3.1	393	13	BJ462915

ALIGNMENTS

RESULT 1  
AQ257253  
LOCUS nbxb0017P08r CUGI Rice BAC Library Oryza sativa genomic clone  
DEFINITION 680 bp DNA linear GSS 23-OCT-1998  
ACCESSION nbxb0017P08r  
VERSION AQ257253.1  
KEYWORDS GSS.  
SOURCE Oryza sativa.  
ORGANISM Oryza sativa  
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
TITLE Ehrhartoideae; Oryzeae; Oryza.  
JOURNAL 1 (bases 1 to 680)  
COMMENT Wing, R.A. and Dean, R.A.  
A BAC End Sequencing Framework to Sequence the Rice Genome  
Unpublished (1998)  
Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
Seq primer: GGAACAGCTATGACCATG  
Class: BAC ends  
High quality sequence start: 47  
High quality sequence stop: 390.  
Location/Qualifiers

```
source
1. .680
/organism="Oryza sativa"
/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/clone="nbx0017P08r"
/clone_lib="CUGI Rice BAC Library"
/tissue_type="Leaf"
/lab_host="E. coli DH10B"
/notes="Vector: pBelOBAC11; Site 1: HindIII; Site 2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."
BASE COUNT 193 a 116 c 127 g 237 t 7 others
ORIGIN
Query Match 10.3%; Score 448; DB 17; Length 680;
Best Local Similarity 93.9%; Pred. No. 4.7e-83;
Matches 477; Conservative 0; Mismatches 30; Indels 1; Gaps 1;
QY 2734 AGCTTGATATGATG-TCTAGATGGTAAGACAATCTCTCGGTGTTCTTAAAAA 2792
Db 41 AAGCTCGAACATGATGTTCTAGATGGTAAGACAATCTCTCGGTGTTCTTAAAAA 100
QY 2793 AAGGTAAGACATACAAATTTTGACATCCCTTTATTTTACTAAATTTTAGTCCACCTGGCT 2852
Db 101 AAGGTAAGACATACAAATTTTGACATCCCTTTATTTTACTAAATTTTAGTCCACCTGGCT 160
QY 2853 GCGGAAAGGGAACACAGTCACCGCTGATTAAGATGAATTTTCTGTGCCATTTAGCCA 2912
Db 161 GCGGAAAGGGAACACAGTCACCGCTGATTAAGATGAATTTTCTGTGCCATTTAGCCA 220
QY 2913 CTGGTGATATGTTGAGGCTGCAGTGGCTGCTAAACTCCACTGGGATTAAGGCTAAG 2972
Db 221 CTGGTGATATGTTGAGGCTGCAGTGGCTGCTAAACTCCACTGGGATTAAGGCTAAG 280
QY 2973 AAGCTATGCAAGGTAGTGTTTTAAAGAAACATATAGCAACAGAAATTAACACGACGA 3032
Db 281 AAGCTATGCAAGGTAGTGTTTTAAAGAAACATATAGCAACAGAAATTAACACGACGA 340
QY 3033 ATGGGTTTCTTGATCTTTTGTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTGATG 3092
Db 341 ATGGGTTTCTTGATCTTTTGTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTGATG 400
QY 3093 ACTTGGTTGTTGGGATTTATGATGAAGCCATGAAGAAATCTTCATGCCAAGGTTTAA 3152
Db 401 ACTTGGTTGTTGGGATTTATGATGAAGCCATGAAGAAATCTTCATGCCAAGGTTTAA 460
QY 3153 TCCTTGATGGTTTCCCTAGAACTGTGTTTCAAGCACAGAAAGGTGAGGTCTTGTGTCAT 3212
Db 461 TCCTTGATGGTTTCCCTAGAACTGTGTTTCAAGCACAGAAAGGTGAGGTCTTGTGTCAT 520
QY 3213 TGCACCGCTATATAAAGAGCTCCCTTT 3240
Db 521 TGCACCTCTACTATATACATTTGATCTT 548
RESULT 2
```

LOCUS AQ289489 520 bp DNA linear GSS 03-DEC-1998  
DEFINITION nbx0035C14f CUGI Rice BAC Library Oryza sativa genomic clone  
ACCESSION nbx0035C14f, DNA sequence.  
VERSION AQ289489  
KEYWORDS AQ289489.1 GI:3950935  
SOURCE GSS.  
ORGANISM Oryza sativa.  
Oryza sativa.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.  
REFERENCE 1 (bases 1 to 520)  
AUTHORS Wing,R.A. and Dean,R.A.  
TITLE A BAC End Sequencing Framework to Sequence the Rice Genome  
JOURNAL Unpublished (1998)  
COMMENT Contact: Wing RA  
Clemson University Genomics Institute  
100 Jordan Hall, Clemson  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
Seq primer: TAATACGACTCACTATAGGG  
Class: BAC ends  
High quality sequence stop: 348.  
Location/Qualifiers  
1. 520  
/organism="Oryza sativa"  
/strain="Japonica"  
/cultivar="Nipponbare"  
/db\_xref="taxon:4530"  
/clone="nbx0035C14f"  
/clone\_lib="CUGI Rice BAC Library"  
/tissue\_type="Leaf"  
/lab\_host="E. coli DH10B"  
/notes="Vector: pBelOBAC11; Site 1: HindIII; Site 2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."

BASE COUNT 148 a 91 c 108 g 173 t

Query Match 10.1%; Score 439.6; DB 17; Length 520;  
Best Local Similarity 94.9%; Pred. No. 2.7e-81;  
Matches 465; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

QY 2751 TCTAGATGGTAAGACAATCTCTCGGTGTTCTTAAAAAAGGTAAGACATACAATT 2810  
Db 6 TCTAGATGGTAAGACAATCTCTCGGTGTTCTTAAAAAAGGTAAGACATACAATT 65  
QY 2811 TTGACATCCCTTATTTTACTAAATTTTAGTCCACCTGGCTGCGGAAAGGGAACACAGT 2870  
Db 66 TTGACATCCCTTATTTTACTAAATTTTAGTCCACCTGGCTGCGGAAAGGGAACACAGT 125  
QY 2871 CACCGCTGATTAAAGGATGAATTTTGTCTTGTCCATTTTAGCCACTGCTGATGTTGAGG 2930  
Db 126 CACCGCTGATTAAAGGATGAATTTTGTCTTGTCCATTTTAGCCACTGCTGATGTTGAGG 185



QY 2931 CTGCAAGGCTGCTAAACCTCCACTTGGGATTAAGCTTAAGAACTATGACAGGTAG 2990  
 |||||  
 Db 186 CTGAGTGGCTGCTAAACCTCCACTTGGGATTAAGCTTAAGAACTATGACAGGTAG 245  
 |||||  
 QY 2991 TTTTAAAGAAACATATAGCAACAGAAATTTATACCAAGCAGGAATGGTTCCTGATTCTT 3050  
 |||||  
 Db 246 -TTTAAAGAAACATATAGCAACAGAAATTTATACCAAGCAGGAATGGTTCCTGATTCTT 304  
 |||||  
 QY 3051 TTGTTCTTCTTCTTCTTCTAGGAGAGCTTCTTCTGATGACTGCTGTGGGATTA 3110  
 |||||  
 Db 305 TTGTTCTTCTTCTTCTTCTAGGAGAGCTTCTTCTGATGACTGCTGTGGGATTA 364  
 |||||  
 QY 3111 TTGATGAAGCCATGAAGAAACTTCATGCCAGAAAGTTTATCTCTGATGGTTCCCTA 3170  
 |||||  
 Db 365 TTGATGAAGCCATGAAGAAACTTCATGCCAGAAAGTTTATCTCTGATGGTTCCCTA 424  
 |||||  
 QY 3171 GAACCTGTGTTTCAGACACAGAGTGGTCCCTTGGTCAATATGACCGCTATATAAG 3230  
 |||||  
 Db 425 AAACCTGGAGTCAATCTCATATATGAGGCTCTTCCTTCATCTCTCGCTATATTTAAA 484  
 |||||  
 QY 3231 AGCTCCTTTT 3240  
 |||||  
 Db 485 TGTCTCTCTT 494  
 |||||

RESULT 3  
 A0915399/c  
 LOCUS nbe0056H06r CUGI Rice BAC Library (EcoRI) linear GSS 02-DEC-1999  
 DEFINITION clone nbe0056H06r, DNA sequence.  
 ACCESSION A0915399  
 VERSION A0915399.1 GI:6511915  
 KEYWORDS GSS.  
 SOURCE Oryza sativa.  
 ORGANISM Oryza sativa.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzeae; Oryza.  
 1 (bases 1 to 502)  
 TITLE A BAC End Sequencing Framework to Sequence the Rice Genome  
 JOURNAL Unpublished (1998)  
 COMMENT Contact: Wing RA  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Tel: 864 656 7288  
 Fax: 864 656 4293  
 Email: rwing@clemson.edu  
 Seq primer: GGAACAGCTATGACCATG  
 Class: BAC ends  
 High quality sequence start: 28  
 High quality sequence stop: 200.  
 Location/Qualifiers  
 1..502  
 /organism="Oryza sativa"  
 /strain="Japonica"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:4530"  
 /clone="nbe0056H06r"  
 /tissue\_type="Leaf"  
 /lab\_host="E. coli DH10B"  
 /notes="Vector: pBACIndigo; Site 1: EcoRI; Site 2: EcoRI;  
 Rice is the most important food crop in the world. Half of  
 the world population, especially those inhabiting highly  
 populated areas of the humid tropics and subtropics, rely  
 on rice as their primary source of carbohydrate.  
 Monocotyledonous rice is a diploid plant (2n=24) with a  
 haploid genome equivalent of 431 Mbp (Arumuganathan and  
 Earle, 1991). The relatively small genome of rice, three  
 times larger than that of Arabidopsis, makes it suitable

for genomic studies. In order to facilitate positional  
 cloning, physical mapping and genome sequencing of rice,  
 we have constructed a BAC library from Oryza sativa,  
 Nipponbare variety using EcoRI as the cloning enzyme. The  
 library contains 55,296 clones with an average insert size  
 of 121 Kb providing approximately 15 haploid genome  
 equivalents. The deep coverage allows the isolation a  
 particular sequence with a probability of 99.9 %. Three  
 high density filters, each containing 18,432 clones  
 (doubly spotted), represent the whole library for colony  
 screening and can be requested from the Clemson University  
 BAC/EST Resource Center (www.genome.clemson.edu)."

BASE COUNT 182 a 97 c 99 g 123 t 1 others  
 ORIGIN  
 Query Match 5.5%; Score 238; DB 17; Length 502;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-39;  
 Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4104 ATTGCTTTGCTTTTTCACAAAGTTTACTGGAGAACCTTAATTCAAAGGAAGATGACAC 4163  
 |||||  
 Db 501 ATTGCTTTGCTTTTTCACAAAGTTTACTGGAGAACCTTAATTCAAAGGAAGATGACAC 442  
 |||||  
 QY 4164 AGCTGCAGTATTGAAGTCAAGGCTTGAAGCCTTCCACGTACAAACTAAGCCTGTATGTTT 4223  
 |||||  
 Db 441 AGCTGCAGTATTGAAGTCAAGGCTTGAAGCCTTCCACGTACAAACTAAGCCTGTATGTTT 382  
 |||||  
 QY 4224 CCTTAGCACTACGTTTTTAAATATTCAGATATTCCTTTTAGGATGTAGTCGACTTCAG 4283  
 |||||  
 Db 381 CCTTAGCACTACGTTTTTAAATATTCAGATATTCCTTTTAGGATGTAGTCGACTTCAG 322  
 |||||  
 QY 4284 TTAAGCGGATTCCTTCAGTTGCATTACAGTGTTCCTGTATATCTTTCATTGTGTTTTT 4341  
 |||||  
 Db 321 TTAAGCGGATTCCTTCAGTTGCATTACAGTGTTCCTGTATATCTTTCATTGTGTTTTT 264  
 |||||

RESULT 4  
 AUI64649  
 LOCUS AUI64649 Rice root Oryza sativa linear EST 03-APR-2002  
 DEFINITION Clone R0936, mRNA sequence.  
 ACCESSION AUI64649  
 VERSION AUI64649.1 GI:11173174  
 KEYWORDS EST.  
 SOURCE Oryza sativa (japonica cultivar-group).  
 ORGANISM Oryza sativa (japonica cultivar-group).  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzeae; Oryza.  
 1 (bases 1 to 287)  
 TITLE Rice cDNA from root (2000)  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Takuji Sasaki  
 National Institute of Agrobiological Resources  
 Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki  
 305-8602, Japan  
 Tel: 81-298-38-7441  
 Fax: 81-298-38-7468  
 Email: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/  
 PROJECT = RGP.  
 Location/Qualifiers  
 1..287  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="R0936"  
 /clone\_lib="Rice root"  
 /note="Prepared from seedling root." 2 others  
 BASE COUNT 91 a 52 c 60 g 82 t 2 others  
 ORIGIN  
 Query Match 4.0%; Score 173.2; DB 9; Length 287;



```

Best Local Similarity 91.9%; Pred. No. 9.5e-21;
Matches 158; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 3312 AGCTTGATGAATGTCGCCAAACAAGGTACTAAGATTGACAAGGTTCTAAATTTGCAA 3371
|||||
Db 441 AGCTGATGAATGTCGCCAAACAAGGTCTAAGGTTGACAAGGTTCTAAATTTGCAA 500
|||||

QY 3372 TTGATGATGCAATCTCGGAGACGATACCGTCTGGATCCACCATCAAGTGGTA 3431
|||||
Db 501 TTGATGATGCAATCTCGGAGACGATACCGTCTGGATCCACCATCAAGTGGTA 560
|||||

QY 3432 GATCTTATCATACAAAATTTGCTCTCTAAGACTCTCGGACTTGATGATGT 3483
|||||
Db 561 GATCTTATCATACAAAATTTGCTCTCTAAGACTCTCGGACTTGATGATGT 612
|||||

RESULT 7
BE493591 BE493591 680 bp mRNA linear EST 16-APR-2001
LOCUS WHE0568_B05_C10ZE Triticum monococcum vegetative apex cDNA library
DEFINITION Triticum monococcum cDNA clone WHE0568_B05_C10, mRNA sequence.
ACCESSION BE493591
VERSION BE493591.1 GI:9660184
KEYWORDS EST.
SOURCE Triticum monococcum.
ORGANISM Triticum monococcum.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 680)
AUTHORS Anderson,O.D., Chao,S., Dubcovsky,J., Echenique,V., Han,P.S., Hsia
,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J., Seaton,C.L.,
Stamove,B. and Tong,J.C.
The structure and function of the expressed portion of the wheat
genomes - vegetative apex cDNA library from Triticum monococcum
Unpublished (2001)
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Stratagene T3 primer.
Location/Qualifiers
1 . 680
/organism="Triticum monococcum"
/cultivar="DV92"
/db_xref="taxon:4568"
/clone_lib="WHE0568_B05_C10"
/library="Triticum monococcum vegetative apex cDNA
library"
/tissue_type="Vegetative shoot apex"
/dev_stage="Three weeks-old plants"
/lab_host="E. coli XL0LR"
/notes="Vector: Lambda pbk-CMV (Lambda Zap Express),
excised phagemid: Site 1: EcoRI; Site 2: XhoI. The tissue,
total RNA, and poly(A)+RNA were prepared, a cDNA library
was made, and the cDNA clones were in vivo excised at the
University of California, Davis (V. Echenique, B. Stamova
, J. Dubcovsky). Plasmid DNA preparations and DNA
sequencing were performed in the OD Anderson lab (all
other authors)."
```

BASE COUNT 202 a 134 c 154 g 190 t

ORIGIN

Query Match 3.4%; Score 149.6; DB 10; Length 680;  
 Best Local Similarity 91.9%; Pred. No. 9.5e-21;  
 Matches 159; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 3312 AGCTTGATGAATGTCGCCAAACAAGGTACTAAGATTGACAAGGTTCTAAATTTGCAA 3371

```

Db 160 AGCTGATGAATGTCGCCAAACAAGGTGCTAAGTTTGACAAGGTTCTAAATTTGCAA 219
|||||
QY 3372 TTGATGATGCAATCTCGGAGACGATACCGTCTGGATCCACCATCAAGTGGTA 3431
|||||
Db 220 TTGATGATGCAATCTCGGAGACGATACCGTCTGGATCCACCATCAAGTGGTA 279
|||||

QY 3432 GATCTTATCATACAAAATTTGCTCTCTAAGACTCTCGGACTTGATGATGT 3483
|||||
Db 280 GATCTTATCATACAAAATTTGCTCTCTAAGACTCTCGGACTTGATGATGT 331
|||||

RESULT 8
BE405815 BE405815 514 bp mRNA linear EST 21-JUL-2000
LOCUS WHE0437_C12_E23ZS Wheat etiolated seedling root cDNA library
DEFINITION Triticum aestivum cDNA clone WHE0437_C12_E23, mRNA sequence.
ACCESSION BE405815
VERSION BE405815.1 GI:9365283
KEYWORDS EST.
SOURCE Bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 514)
AUTHORS Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J.,
Seaton,C.L. and Tong,J.C.
The structure and function of the expressed portion of the wheat
genomes
Unpublished (2000)
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.
Location/Qualifiers
1 . 514
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone_lib="WHE0437_C12_E23"
/clone_lib="Wheat etiolated seedling root cDNA library"
/tissue_type="Root"
/dev_stage="Five day old etiolated seedling"
/lab_host="E. coli SOLR"
/notes="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site 1: EcoRI; Site 2: XhoI. Seeds were surface-sterilized
, germinated and grown aseptically in the dark at room
temperature on filter paper with water, nystatin and
cefotaxime in covered crystallization dishes. Roots were
harvested. The tissue, total RNA, and poly(A) RNA were
prepared, a cDNA library was made, and the cDNA clones
were in vivo excised to give pBluescript phagemids in the
TJ Close lab (Choi, Close, Fenton) at the University of
California, Riverside. Plasmid DNA preparations and DNA
sequencing were performed in the OD Anderson lab (all
other authors)."
```

BASE COUNT 136 a 115 c 128 g 134 t 1 others

ORIGIN

Query Match 3.4%; Score 148.6; DB 10; Length 514;  
 Best Local Similarity 91.8%; Pred. No. 1.5e-20;  
 Matches 157; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 3312 AGCTTGATGAATGTCGCCAAACAAGGTACTAAGATTGACAAGGTTCTAAATTTGCAA 3371

```

Db 344 AGCTGATCAATGCTGGCAAGCAAGTGCTGAAGTTTGACAGGTTCTAAATTTTGCAA 403
QY 3372 TTGATGATGCATCTGGAAGCAAGTAATACCGTCTGTGGATCCACCATCAAGTGGA 3431
Db 404 TTGATGATGCATCTGGAAGCAAGTAATACCGTCTGTGGATCCACCATCAAGTGGA 463
QY 3432 GATCTTATCATCAAAATTTGCTCTCTCTTAAGACTCCTGGACTTGGATGATG 3482
Db 464 GATCTTATCATCAAAATTTGCTCTCTCTTAAGACTCCTGGACTTGGATGATG 514

RESULT 9
BQ840903 774 bp mRNA linear EST 12-AUG-2002
LOCUS WHE4204_G10_N20Zs Aegilops speltoides pre-meiotic anther cDNA
DEFINITION library Aegilops speltoides cDNA clone WHE4204_G10_N20, mRNA
sequence.
ACCESSION BQ840903
VERSION BQ840903.1 GI:22210312
KEYWORDS EST:
SOURCE Aegilops speltoides.
ORGANISM Aegilops speltoides.
REFERENCE 1. (bases 1 to 774)
AUTHORS Akhunov,E., Anderson,O.D., Chao,S., Crossman,C., Devorak,J., Lazo
,G.R., Pham,J., Rauch,C.J., Wilson,C. and Woo,J.
TITLE The structure and function of the expressed portion of the wheat
genomes - Pre-meiotic anther cDNA library from Aegilops speltoides
JOURNAL Unpublished (2002)
COMMENT Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: SK primer.

FEATURES
source
1..774
/organism="Aegilops speltoides"
/cultivar="P2 from 2-12-4-8-1-1-1-1" x PI36909-12-811-1(1)
/db_xref="taxon:4573"
/clone="WHE4204_G10_N20"
/clone_lib="Aegilops speltoides pre-meiotic anther cDNA
library"
/tissue_type="Anther"
/dev_stage="Pre-meiotic anthers"
/lab_host="E. coli SOLR"
/notes="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site 1: EcoRI; Site 2: XhoI; Plants were grown in a growth
chamber at the University of California, Davis (Akhunov in
Devorak's lab). Pre-meiotic anthers were harvested, total
RNA and poly(A) RNA were prepared, a cDNA library was made
, and the cDNA clones were in vivo excised to give
pBluescript phagemids at the University of California,
Davis (Akhunov in Devorak's lab). Plasmid DNA preparations
and DNA sequencing were performed in the OD Anderson lab
(all other authors)."
BASE COUNT 209 a 176 c 204 g 185 t
ORIGIN
Query Match 3.4%; Score 148; DB 14; Length 774;
Best Local Similarity 91.3%; Pred. No. 2.1e-20;
Matches 157; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 3312 AGCTTGATGAATGTGGCCAAACAGGTACTAGATTGACAGGTTCTAAATTTTGCAA 3371
Db 467 AGCTGATGAATGTGGCCAAACAGGTACTAGATTGACAGGTTCTAAATTTTGCAA 526

```

```

QY 3372 TTGATGATGCATCTGGAAGCAAGTAATACCGTCTGTGGATCCACCATCAAGTGGA 3431
Db 527 TTGATGATGCATCTGGAAGCAAGTAATACCGTCTGTGGATCCACCATCAAGTGGA 586
QY 3432 GATCTTATCATCAAAATTTGCTCTCTCTTAAGACTCCTGGACTTGGATGATG 3483
Db 587 GATCTTATCATCAAAATTTGCTCTCTCTTAAGACTCCTGGACTTGGATGATG 638

RESULT 10
BQ443438 596 bp mRNA linear EST 25-JUL-2000
LOCUS WHE1104_B09_D18Zs Wheat etiolated seedling root normalized cDNA
DEFINITION library Triticum aestivum cDNA clone WHE1104_B09_D18, mRNA
sequence.
ACCESSION BQ443438
VERSION BQ443438.1 GI:9442972
KEYWORDS EST:
SOURCE Bread wheat.
ORGANISM Triticum aestivum
REFERENCE 1. (bases 1 to 596)
AUTHORS Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
,P.S., Haia,C.C., Kang,Y., Lazo,G.R., Miller,R., Nguyen,H.T.,
Rauch,C.J., Seaton,C.L., Tong,J.C. and Zhang,B.
TITLE The structure and function of the expressed portion of the wheat
genomes - Normalized root cDNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.

FEATURES
source
1..596
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE1104_B09_D18"
/clone_lib="Wheat etiolated seedling root normalized cDNA
library"
/tissue_type="Root"
/dev_stage="Five day old etiolated seedling"
/lab_host="E. coli DH10B"
/notes="Vector: Lambda Uni-ZAP XR, excised phagemid
pBluescript SK; Site 1: EcoRI; Site 2: XhoI; Seeds were
surface-sterilized, germinated and grown aseptically in
the dark at room temperature on filter paper with water,
nystatin and cefotaxime in covered crystallization
dishes. Roots were harvested. The tissue, total RNA, and
poly(A) RNA were prepared, a cDNA library was made in the
TJ Close lab (Choi, Close, Fenton) at the University of
California, Riverside. The cDNA clones were in vivo
excised to give pBluescript phagemids before
normalization was carried out. The mass excision of
phagemid library and normalization were done in HT Nguyen
lab by D. Zhang at Texas Tech University. Normalization
protocol used was that of Soares. Plasmid DNA
preparations and DNA sequencing were performed in the OD
Anderson lab (all other authors)."
BASE COUNT 157 a 138 c 156 g 145 t
ORIGIN
Query Match 3.4%; Score 147.6; DB 10; Length 596;
Best Local Similarity 91.8%; Pred. No. 2.5e-20;

```

```

Matches 156; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 3312 AGCTTGATGAATGTTGGCCAAACAGGTTACTAAGATTGACAAGTTCTAAATTTTGCAA 3371
Db 427 AGCTCGATGAATGCTGGCAAGCAGGTTGCTAAGGTTGACAAGGTTCTAAATTTTGCAA 486
Qy 3372 TTGATGATGCAATCTGGAAGAACGAATTAACCGGTCGTTGGATCCACCCATCAAGTGGTA 3431
Db 487 TTGATGATGCAATATGGAAGACGAATTAACCGGTCGTTGGATCCACCCATCAAGTGGTA 546
Qy 3432 GATCTTATCATACAAAATTTGCTCTCTTAAGACTCTCGACTTGATGAT 3481
Db 547 GATCTTACCATACAAAATTTGCTCTCTTAAGACTCCAGGAGTTGATGAT 596

RESULT 11
LOCUS BQ243888/c 612 bp mRNA linear EST 03-MAY-2002
DEFINITION TAE15007G03F TAE15 Triticum aestivum cDNA clone TAE15007G03F, mRNA
sequence.
ACCESSION BQ243888
VERSION BQ243888.1 GI:20439764
KEYWORDS EST.
SOURCE Bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 612)
AUTHORS Cloutier,S.
JOURNAL Wheat functional genomics - Glenlea developing seeds cDNA libraries
COMMENT Unpublished (2002)
Contact: Dr. Sylvie Cloutier
Cereal Research Centre, Agriculture and Agri-food Canada
195 Dafoe Rd, Winnipeg, MB, Canada R3T 2M9
Tel: (204) 983-2340
Fax: (204) 983-4604
Email: scloutier@agr.ca
was cloned directionally, not all sequences generated with reverse
primer were from the 5' end (same with forward primer and 3' end).
Average insert size is >1.4 kb
Plate: 007 row: G column: 03
Seq primer: M13 Forward.
FEATURES
source
1..612
/organism="Triticum aestivum"
/cultivar="Glenlea"
/db_xref="taxon:4565"
/clone="TAE15007G03F"
/tissue_type="developing seeds"
/dev_stage="15 days after anthesis"
/lab_host="E. coli DH10B"
/notes="Vector: pCMV-SPORT6.0 (Invitrogen Technologies);
Site 1: NotI; Site 2: MluI; mRNA obtained from wheat seeds
of cultivar Glenlea 15 days post-anthesis"
BASE COUNT 180 a 131 c 125 g 176 t
ORIGIN
1..612
/organism="Triticum aestivum"
/cultivar="Glenlea"
/db_xref="taxon:4565"
/clone="TAE15007G03F"
/tissue_type="developing seeds"
/dev_stage="15 days after anthesis"
/lab_host="E. coli DH10B"
/notes="Vector: pCMV-SPORT6.0 (Invitrogen Technologies);
Site 1: NotI; Site 2: MluI; mRNA obtained from wheat seeds
of cultivar Glenlea 15 days post-anthesis"
Query Match 3.4%; Score 146.4; DB 14; Length 612;
Best Local Similarity 90.7%; Pred. No. 4.4e-20;
Matches 156; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 3312 AGCTTGATGAATGTTGGCCAAACAGGTTACTAAGATTGACAAGTTCTAAATTTTGCAA 3371
Db 555 AGCTCGATGAATGCTGGCAAGCAGGTTGCTAAGGTTGACAAGTTCTAAATTTTGCAA 496
Qy 3372 TTGATGATGCAATCTGGAAGAACGAATTAACCGGTCGTTGGATCCACCCATCAAGTGGTA 3431
Db 495 TTGATGATGCAATATGGAAGAACGAATTAACCGGTCGTTGGATCCACCCATCAAGTGGTA 436
Qy 3432 GATCTTATCATACAAAATTTGCTCTCTTAAGACTCTCGACTTGATGAT 3483

```

```

Db 435 GATCTTACCATACAAAATTTGCTCTCTCCCAAGACTCCGGGAGTTGATGATGT 384

RESULT 12
LOCUS BQ246627 719 bp mRNA linear EST 03-MAY-2002
DEFINITION TAE15007G03R TAE15 Triticum aestivum cDNA clone TAE15007G03R, mRNA
sequence.
ACCESSION BQ246627
VERSION BQ246627.1 GI:20442503
KEYWORDS EST.
SOURCE Bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 719)
AUTHORS Cloutier,S.
JOURNAL Wheat functional genomics - Glenlea developing seeds cDNA libraries
COMMENT Unpublished (2002)
Contact: Dr. Sylvie Cloutier
Cereal Research Centre, Agriculture and Agri-food Canada
195 Dafoe Rd, Winnipeg, MB, Canada R3T 2M9
Tel: (204) 983-2340
Fax: (204) 983-4604
Email: scloutier@agr.ca
was cloned directionally, not all sequences generated with reverse
primer were from the 5' end (same with forward primer and 3' end).
Average insert size is >1.4 kb
Plate: 007 row: G column: 03
Seq primer: M13 Reverse.
FEATURES
source
1..719
/organism="Triticum aestivum"
/cultivar="Glenlea"
/db_xref="taxon:4565"
/clone="TAE15007G03R"
/tissue_type="developing seeds"
/dev_stage="15 days after anthesis"
/lab_host="E. coli DH10B"
/notes="Vector: pCMV-SPORT6.0 (Invitrogen Technologies);
Site 1: NotI; Site 2: MluI; mRNA obtained from wheat seeds
of cultivar Glenlea 15 days post-anthesis"
BASE COUNT 207 a 148 c 163 g 201 t
ORIGIN
1..719
/organism="Triticum aestivum"
/cultivar="Glenlea"
/db_xref="taxon:4565"
/clone="TAE15007G03R"
/tissue_type="developing seeds"
/dev_stage="15 days after anthesis"
/lab_host="E. coli DH10B"
/notes="Vector: pCMV-SPORT6.0 (Invitrogen Technologies);
Site 1: NotI; Site 2: MluI; mRNA obtained from wheat seeds
of cultivar Glenlea 15 days post-anthesis"
Query Match 3.4%; Score 146.4; DB 14; Length 719;
Best Local Similarity 90.7%; Pred. No. 4.4e-20;
Matches 156; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 3312 AGCTTGATGAATGTTGGCCAAACAGGTTACTAAGATTGACAAGTTCTAAATTTTGCAA 3371
Db 230 AGCTCGATGAATGCTGGCAAGCAGGTTGCTAAGGTTGACAAGTTCTAAATTTTGCAA 289
Qy 3372 TTGATGATGCAATCTGGAAGAACGAATTAACCGGTCGTTGGATCCACCCATCAAGTGGTA 3431
Db 290 TTGATGATGCAATATGGAAGAACGAATTAACCGGTCGTTGGATCCACCCATCAAGTGGTA 349
Qy 3432 GATCTTATCATACAAAATTTGCTCTCTTAAGACTCTCGACTTGATGATGT 3483
Db 350 GATCTTACCATACAAAATTTGCTCTCTCCCAAGACTCCGGGAGTTGATGATGT 401

RESULT 13
LOCUS BG418644 864 bp mRNA linear EST 23-OCT-2001
DEFINITION HVSMEK0023K16f Hordeum vulgare testae pericarp EST library
HVGDNA0013 (normal) Hordeum vulgare cDNA clone HVSMEK0023K16f, mRNA
sequence.
ACCESSION BG418644
VERSION BG418644.1 GI:13324195
KEYWORDS EST.

```

```

SOURCE      Hordeum vulgare.
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1 (bases 1 to 864)
AUTHORS    Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Kannangara,G., von
            Wetstein,D., Akhunov,E., Chin,A., Choi,D.W., Fenton,R.D., Kianian
            ,P., Otto,C., Simons,K., Zhang,D., Begum,D., Begum,D., Yu,Y.,
            Henry,D., Palmer,M., Rambo,T., Simmons,J., Oates,R. and Main,D.
TITLE      Development of a genetically and physically anchored EST resource
            for barley genomics: Morex testa/pericarp cDNA library
JOURNAL    Unpublished (2001)
COMMENT    Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Total hg bases=535
            Seq primer: AATTACCCCTCACTAAGGG
            High quality sequence stop: 609.
            Location/Qualifiers
                1..864
                /organism="Hordeum vulgare"
                /cultivar="Morex"
                /db_xref="taxon:4513"
                /clone="HVSMEK0023K16f"
                /clone_lib="Hordeum vulgare testa/pericarp EST library
                HVCNDA0013 (normal)"
                /tissue types="testa/pericarp"
                /lab host="TJC121"
                /note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;
                Plants were raised from seeds in a Controlled Environments
                growth chamber maintained in continuous light at 18°C, and
                testa and pericarp were dissected from developing kernels
                at Washington State University, Pullman, WA (Kannangara,
                von Wetstein). Total RNA was prepared, poly(A) RNA was
                purified, one cDNA library was made, and 1 million pfu
                were in vivo excised to give pBluescript SK(-) cDNA
                phagemids in the TJ Close lab at the University of
                California, Riverside (Akhunov, Chin, Choi, Close, Fenton,
                Kianian, Otto, Simons, Zhang). Phagemids were plated and
                picked at the Clemson University Genomics Institute (CUGI)
                (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA
                preparations, DNA sequencing and sequence analysis were
                performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates
                , Rambo, Main). The sequence has been trimmed to remove
                vector sequence and contains a minimum of 100 bases of
                phred value 20 or above. For more details on library
                preparation and sequence analysis see
                http://www.genome.clemson.edu/projects/barley. To order
                this clone see http://www.genome.clemson.edu/orders
                see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
                Genetically and physically anchored EST resources for
                barley genomics. Barley Genetics Newsletter 31:29-30.
                (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"
BASE COUNT  220 a 175 c 178 g 291 t
ORIGIN
Query Match      3.3%; Score 145.2; DB 12; Length 864;
Best Local Similarity 89.7%; Pred. No. 7.9e-20;
Matches 156; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy  3310 GCAGCTTCATGAATTTGGCCAAACAAGGTACTAAGATTGACAGGTTCTTAATTTTGC 3369
      |||||
Db    1 GCAGCTTCATGAATGCTGGCAAGCAAGGTGCTAAGTTGACAGGTTCTGAATTTGC 60
      |||||
Qy  3370 AATTGATGATCAATCTGGAAGAACGAATTACCGGTGCTTGGATCCACCCATCAAGTGG 3429
      |||||
Db    61 AATTGATGATCAATATTGGAAGAACGAATTACTGGCCGTTTGATACACCCATCGAGTGG 120
      |||||

Qy  3430 TAGATCTTATCATACAAAATTTGCTCTCTTAAGACTCCTGGACTTGTATGATGT 3483
      |||||
Db    121 TAGATCTTACCATACAAAATTTGCTCTCTTAAGACTCCGGAGTTGTATGATGT 174
      |||||

RESULT 14
BE495096/c
LOCUS      BE495096
DEFINITION WHE1274_H05_P10Z5 Secale cereale anther cDNA library Secale cereale
            cDNA clone WHE1274_H05_P10, mRNA sequence.
ACCESSION BE495096
VERSION   BE495096.1 GI:9661689
KEYWORDS  EST.
SOURCE    rye.
ORGANISM  Secale cereale
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Secale.
REFERENCE  1 (bases 1 to 593)
AUTHORS    Anderson,O.D., Butler,B., Chao,S., Choi,D.W., Close,T.J., Fenton
            ,R.D., Gustafson,J.P., Han,P.S., Hsia,C.C., Kang,Y., Lazo,G.R.,
            Miller,R., Rausch,C.J., Ross,K., Seaton,C.L. and Tong,J.C.
TITLE      The structure and function of the expressed portion of the wheat
            genomes - Anther cDNA library from rye
JOURNAL    Unpublished (2000)
COMMENT    Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oanderson@pw.usda.gov
            Sequence have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: Stratagene SK primer.
            Location/Qualifiers
                1..593
                /organism="Secale cereale"
                /cultivar="Blanco"
                /db_xref="taxon:4550"
                /clone="WHE1274_H05_P10"
                /clone_lib="Secale Cereale anther cDNA library"
                /tissue type="Anther"
                /dev stage="Adult plant before anthesis"
                /lab host="E. coli SOLR"
                /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                Site 1: EcoRI; Site 2: XhoI; Plants were grown in the
                greenhouse. Anthers were harvested and pooled from early
                meiosis to late meiosis. The tissue, total RNA, and
                poly(A) RNA were prepared (Butler, Ross and Gustafson) at
                University of Missouri, Columbia. A cDNA library was
                made, and the cDNA clones were in vivo excised to give
                phagescript phagemids in the TJ Close lab (Choi, Close,
                Fenton) at the University of California, Riverside.
                Plasmid DNA preparations and DNA sequencing were performed
                in the OD Anderson lab (all other authors)."
BASE COUNT  165 a 125 c 114 g 189 t
ORIGIN
Query Match      3.3%; Score 144.8; DB 10; Length 593;
Best Local Similarity 90.1%; Pred. No. 9.6e-20;
Matches 155; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy  3312 AGCTTGATGAATTTGGCCAAACAAGGTACTAAGATTGACAGGTTCTTAATTTTGC 3371
      |||||
Db    579 AGCTCGATGAATGCTGGCAAGCAAGGTGCTAAGTTGACAGGTTCTGAATTTTGC 520
      |||||
Qy  3372 TTGATGATGAATCTGGAAGACGAATTACCGGTGCTTGGATCCACCCATCAAGTGT 3431
      |||||
Db    519 TTGATGATGAATATTGGAAGAACCGGATCAGTGGCCGTTGGATACACCCATCGAGTGT 460
      |||||
Qy  3432 GATCTTATCATACAAAATTTGCTCTCTTAAGACTCCTGGACTTGTATGATGT 3483
      |||||

```

Db 459 GATCTTACCATACAAAATTGCTCTCTTAAGACTCCGGAGTTGATGATGT 408

## RESULT 15

BE403629

LOCUS

DEFINITION

BE403629 BE403629 639 bp mRNA linear EST 21-JUL-2000

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae

; Triticeae; Triticum.

1 (bases 1 to 639)

Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D., Han

, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J.,

Seaton, C.L. and Tong, J.C.

The structure and function of the expressed portion of the wheat

genomes

Unpublished (2000)

Contact: Olin Anderson

US Department of Agriculture, Agriculture Research Service, Pacific

West Area, Western Regional Research Center

800 Buchanan Street, Albany, CA 94710, USA

Tel: 5105595773

Fax: 5105595818

Email: oanders@pw.usda.gov

Sequence have been trimmed to remove vector sequence and low

quality sequence with phred score less than 20

Seq Primer: Strategene SK primer.

Location/Qualifiers

1..639

/organism="Triticum aestivum"

/cultivar="Chinese Spring"

/db\_xref="taxon:4565"

/clone="WHE0434 H01\_002"

/clone\_lib="Wheat etiolated seedling root cDNA library"

/tissue\_type="Root"

/dev\_stage="Five day old etiolated seedling"

/lab\_host="E. coli SOLR"

/note="Vector: Lambda Uni-ZAP XR, excised phagemid;

Site\_1: EcoRI; Site\_2: XhoI; Seeds were surface-sterilized

, germinated and grown aseptically in the dark at room

temperature on filter paper with water, nystatin and

cefotaxime in covered crystallization dishes. Roots were

harvested. The tissue, total RNA, and poly(A) RNA were

prepared. A cDNA library was made, and the cDNA clones

were in vivo excised to give phagescript phagemids in the

TJ Close lab (Choi, Close, Fenton) at the University of

California, Riverside. Plasmid DNA preparations and DNA

sequencing were performed in the OD Anderson lab (all

other authors)."

BASE COUNT 167 a 148 c 170 g 154 t

ORIGIN

Query Match 3.3%; Score 144.8; DB 10; Length 639;

Best Local Similarity 90.1%; Pred. No. 9.6e-20;

Matches 155; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 3312 AGCTTGATGAAATGTTGGCCAAACAGGTACTTAAGATTGACAAAGGTTCTTAATTTTGCAA 3371

DB 441 AGCTCGATGAAATGTTGGCAAGCAAGGTCTAAGGTTGGCAAGGTTCTTAATTTTGCAA 500

QY 3372 TTGATGATGCAATACTGGAAGACGAATTACCGTCTGTTGGATCCACCATCAAGTGTA 3431

DB 501 TTGATGATGCAATATTGGAAGACGAATTACTGGCCGTTGGATACACCATCGAGTGTA 560

QY 3432 GATCTTATCATACAAAATTTGCTCTCTTAAGACTCCTCGACTTGATGATGT 3483

DB 561 GATCTTATCATACAAAATTTGCTCTCTCTCAAGACTCCGGAGTTGATGATGT 612

Search completed: March 15, 2003, 16:52:38  
Job time : 5386.47 secs





GenCore version 5.1.4.p5.4578  
Copyright (C) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2003, 09:16:40 ; Search time 10726.7 Seconds  
(without alignments)  
11777.621 Million cell updates/sec

Title: US-09-802-937-1  
Perfect score: 4341  
Sequence: 1 ctgcaggagattaattagg.....gtatatcttcattgtttt 4341

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- 1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*  
29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pin.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
2	3976.8	91.6	98782	2	CNS08C85	AL731756 Oryza sat
C 3	277.8	6.4	58527	8	AB007649	AB007649 Arabidops
C 4	229	5.3	1154	8	RICADXA	D10334 Oryza sativ
C 5	171.2	3.9	8015	8	AB012248	AB012248 Arabidops
C 6	164.4	3.8	387	8	RICKN57	D10429 Rice mrna f
C 7	118.8	2.7	1085	8	RICADXB	D10335 Oryza sativ
C 8	115.6	2.7	13748	8	AP003297	AP003297 Oryza sat
C 9	112.8	2.6	772	8	AY079340	AY079340 Arabidops
C 10	112.8	2.6	904	8	AY085188	AY085188 Arabidops
C 11	112.8	2.6	979	8	AY039945	AY039945 Arabidops
C 12	109.2	2.5	124815	2	AP005594	AP005594 Oryza sat
C 13	109.2	2.5	138025	8	AP003261	AP003261 Oryza sat
C 14	109.2	2.5	169863	2	CNS071PX	AL713952 Oryza sat
C 15	108.8	2.5	126859	8	OSJN00015	AL606455 Oryza sat
C 16	108.8	2.5	155582	2	OSJN00094	AL606649 Oryza sat
C 17	108.4	2.5	100419	2	OSIG00047	AL732355 Oryza sat
C 18	107.6	2.5	100635	8	AP003446	AP003446 Oryza sat
C 19	107.6	2.5	144074	8	AP003302	AP003302 Oryza sat
C 20	106.4	2.5	156054	8	AB023482	AB023482 Oryza sat
C 21	104.8	2.4	129838	8	AC079038	AC079038 Oryza sat
C 22	103.6	2.4	170025	2	OSJN00031	AL606590 Oryza sat
C 23	102.6	2.4	1023	8	AY086140	AY086140 Arabidops
C 24	102	2.3	131983	8	AC023240	AC023240 Oryza sat
C 25	101	2.3	143959	8	AC090055	AC090055 Oryza sat
C 26	100	2.3	6735	8	RIC3H3M	L28995 Oryza sativ
C 27	100	2.3	129845	2	AP004165	AP004165 Oryza sat
C 28	99.4	2.3	121615	2	AP003981	AP003981 Oryza sat
C 29	99.4	2.3	146670	8	AP003250	AP003250 Oryza sat
C 30	99	2.3	142596	8	AP003314	AP003314 Oryza sat
C 31	99	2.3	166753	8	AP002866	AP002866 Oryza sat
C 32	98.4	2.3	159893	8	OSJN00095	OSJN000171
C 33	98.4	2.3	189834	8	CNS08C8F	AL662974 Oryza sat
C 34	96.8	2.2	108805	2	CNS08C89	AL731875 Oryza sat
C 35	96.8	2.2	150238	2	CNS08C89	AL831796 Oryza sat
C 36	96.6	2.2	151100	8	AP003453	AP003453 Oryza sat
C 37	96.2	2.2	136553	2	AC093953	AC093953 Oryza sat
C 38	95.8	2.2	132407	8	OSJN00117	AL607002 Oryza sat
C 39	95.2	2.2	139566	2	AP004161	AP004161 Oryza sat
C 40	95.2	2.2	147954	2	AP003983	AC025296 Oryza sat
C 41	95.2	2.2	165394	8	AC025296	AC025296 Oryza sat
C 42	94.8	2.2	160174	2	AP005103	AP005103 Oryza sat
C 43	94.8	2.2	162010	2	AP004121	AP004121 Oryza sat
C 44	94.6	2.2	102119	2	AP003919	AP003919 Oryza sat
C 45	93.2	2.1	122559	2	AP004301	AP004301 Oryza sat

ALIGNMENTS

RESULT 1  
AB041773  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE

AB041773  
Oryza sativa Adk-a gene for adenylyate kinase, partial cds.  
AB041773.1 GI:7630192  
adenylyate kinase.  
Oryza sativa DNA.  
Oryza sativa  
Eukaryota; Viridiplantae; Sreptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.  
1 (bases 1 to 4341)  
Fukuzawa,H., Uchimiya,H., Tagawa,M. and Arai,S.  
Oryza sativa AK-a gene,promoter region and partial ORF

JOURNAL Published Only in Database (2000)  
REFERENCE 2 (bases 1 to 4341)  
AUTHORS Fukuzawa,H., Uchiyama,H., Tagawa,M. and Arai,S.  
TITLE Direct Submission  
JOURNAL Submitted (12-APR-2000) Hiromitsu Fukuzawa, Nissan Chemical  
Industries,Ltd, Shirooka Research Station of Biological Science,  
1470 Shirooka, Minamisaitama, Saitama 349-0294, Japan  
(E-mail:fukuzawa@nissaichem.co.jp, Tel:+81-480-92-2513)  
Location/Qualifiers

## FEATURES

source 1. .4341  
/organism="Oryza sativa"  
/db\_xref="taxon:4530"  
gene join(1475..1577,2841..2986,3074..3193,3314..3481,  
4126..4215)  
/gene="Adk-a"  
CDS join(1475..1577,2841..2986,3074..3193,3314..3481,  
4126..4215)  
/gene="Adk-a"  
/codon\_start=1  
/product="adenylate kinase"  
/protein\_id="BA094761.1"  
/db\_xref="GI:7630193"  
/translation="MAANLSDVFSMELTTELLRMKSSKPKRVLVGPQCKGKTQ  
SPLKDFCLCHLATQDMLRAVAAKTPGKAKEMDKGELVSDLVVGIIDEAMKK  
TSCOKGFLDGFRTVQAOKLDEMLAKQGTIKDKVLEAIDAILERITGRWIHPS  
SGRSYHFKFAPPTPGUDVDTGELIQRKDDTAAVLKSLEAFHVQTRP"  
BASE COUNT 1284 a 758 c 880 g 1419 t  
ORIGIN

Query Match 100.0%; Score 4341; DB 8; Length 4341;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CTGAGGAGATTAAATAGGTGGACACACCAACCCCTGTGGTGGTGCACCCCTGTGTT	60
DB	1	CTGAGGAGATTAAATAGGTGGACACACCAACCCCTGTGGTGGTGCACCCCTGTGTT	60
QY	61	ATCAACTGGGGTGTGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGG	120
DB	61	ATCAACTGGGGTGTGGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGG	120
QY	121	AGAATGCTCAAGCTGACATGAGAAACGTAATCCAATGGAAGCAATTTCAAGTGGTTC	180
DB	121	AGAATGCTCAAGCTGACATGAGAAACGTAATCCAATGGAAGCAATTTCAAGTGGTTC	180
QY	181	TCTTGTAACCATGTTAGATACATAGACAGTGCACAGTGTGATGGCTCTATATGG	240
DB	181	TCTTGTAACCATGTTAGATACATAGACAGTGCACAGTGTGATGGCTCTATATGG	240
QY	241	CTCGTGTGATCTGACTGTGTGCAAAAGCATCAATTTGCTTGGAGTATCTTTATTA	300
DB	241	CTCGTGTGATCTGACTGTGTGCAAAAGCATCAATTTGCTTGGAGTATCTTTATTA	300
QY	301	CCGAAACCCCAAGATTATTTCTTCACTCAGGTAATTTGCTGAACTATGCAATG	360
DB	301	CCGAAACCCCAAGATTATTTCTTCACTCAGGTAATTTGCTGAACTATGCAATG	360
QY	361	AATACAAATTCGCAAAATATCATGTTATCTATCTCTCAAAATGAAATTTGAGTCCAA	420
DB	361	AATACAAATTCGCAAAATATCATGTTATCTATCTCTCAAAATGAAATTTGAGTCCAA	420
QY	421	CTGAGCTGCAATAGATTTCTTTTCAAAAGAAATTTAATTTTCTTTTCTTTTCTTT	480
DB	421	CTGAGCTGCAATAGATTTCTTTTCAAAAGAAATTTAATTTTCTTTTCTTTTCTTT	480
QY	481	CGCAATTCACCGTTTCGAGAAATATGCTGTCATAAATAGTGTAGTGCAGAAACAA	540
DB	481	CGCAATTCACCGTTTCGAGAAATATGCTGTCATAAATAGTGTAGTGCAGAAACAA	540
QY	541	ATTATATACATAAAAGAGTTTGTATTAATTAACAACCATGTTTCGTACTCAACTCT	600
DB	541	ATTATATACATAAAAGAGTTTGTATTAATTAACAACCATGTTTCGTACTCAACTCT	600

QY	601	AATTTGTAATTTCTTATTTTTCAGTCACAAAATTTCCAATTTCCAATTAAGAAAAATAACCT	660
DB	601	AATTTGTAATTTCTTATTTTTCAGTCACAAAATTTCCAATTTCCAATTAAGAAAAATAACCT	660
QY	661	AGACGGCTAGAGCCACCCATCTAAGGCTAAGTTTCGAGAGGTGAAGTACCGACGAAAAATA	720
DB	661	AGACGGCTAGAGCCACCCATCTAAGGCTAAGTTTCGAGAGGTGAAGTACCGACGAAAAATA	720
QY	721	TGATGGTTTAAATATGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA	780
DB	721	TGATGGTTTAAATATGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA	780
QY	781	TATCATTTAATGTTTGAAGAACGTGCACATAAGAAAACTAAGAACGATGAGTTGGAAAA	840
DB	781	TATCATTTAATGTTTGAAGAACGTGCACATAAGAAAACTAAGAACGATGAGTTGGAAAA	840
QY	841	CAAGAGAAAAACACAGCTTTAAGGCTTCTTGAATCTTCTAGTCTTCTAGTCTTCTTCTT	900
DB	841	CAAGAGAAAAACACAGCTTTAAGGCTTCTTGAATCTTCTAGTCTTCTAGTCTTCTTCTT	900
QY	901	GCATGATAAACGAGAAAGCTCATTAGACATTAATTAATTAATTAATTAATTAATTAATTA	960
DB	901	GCATGATAAACGAGAAAGCTCATTAGACATTAATTAATTAATTAATTAATTAATTAATTA	960
QY	961	TGAAAAAATATTTATTTGAATTTTAAACAATGATGATGATGATGATGATGATGATGATG	1020
DB	961	TGAAAAAATATTTATTTGAATTTTAAACAATGATGATGATGATGATGATGATGATGATG	1020
QY	1021	ACACCAATTTAAACCTTTAAAGAGCATCTTAATAGAAAAAGGAGGAAATTAAGATTCAC	1080
DB	1021	ACACCAATTTAAACCTTTAAAGAGCATCTTAATAGAAAAAGGAGGAAATTAAGATTCAC	1080
QY	1081	GAAGTGTTCGATAATGAAAAATGCGGTGGGATTGAATTTGGTAAATGAATCAGGTTAG	1140
DB	1081	GAAGTGTTCGATAATGAAAAATGCGGTGGGATTGAATTTGGTAAATGAATCAGGTTAG	1140
QY	1141	GATTAATATTTAAATGAAAGAGGAGAAATGAATTTGTTAGATTTAAATGTTCTTTT	1200
DB	1141	GATTAATATTTAAATGAAAGAGGAGAAATGAATTTGTTAGATTTAAATGTTCTTTT	1200
QY	1201	GTGGGTAGAAAAATTTATTTGCTATACACTCCCGAGAGCGTGTGCTTGGTGGGAG	1260
DB	1201	GTGGGTAGAAAAATTTATTTGCTATACACTCCCGAGAGCGTGTGCTTGGTGGGAG	1260
QY	1261	AGCTCTTTTCTGTTGAAAAAATCTGTTTAAAGAAAAAGGAGGAGGAGGAGGAGGAG	1320
DB	1261	AGCTCTTTTCTGTTGAAAAAATCTGTTTAAAGAAAAAGGAGGAGGAGGAGGAGGAG	1320
QY	1321	GTTGTCAACCTCTCACAGAAACGAAACAAAGAGGAGGAGGAGGAGGAGGAGGAGGAG	1380
DB	1321	GTTGTCAACCTCTCACAGAAACGAAACAAAGAGGAGGAGGAGGAGGAGGAGGAGGAG	1380
QY	1381	GACGAGACTCTCCATTTTACGCGGCGACCGGAGAGGAGGAGGAGGAGGAGGAGGAG	1440
DB	1381	GACGAGACTCTCCATTTTACGCGGCGACCGGAGAGGAGGAGGAGGAGGAGGAGGAG	1440
QY	1441	GAAACAGGCTAGTGCAGTGTGCTGCGGAGATGCGGCGGAGGAGGAGGAGGAGGAG	1500
DB	1441	GAAACAGGCTAGTGCAGTGTGCTGCGGAGATGCGGCGGAGGAGGAGGAGGAGGAG	1500
QY	1501	GTCGATGGAGCTGATGACGAGGCTGCTCCGCGCATGAAGTGCAGTCCAAAGCCGACA	1560
DB	1501	GTCGATGGAGCTGATGACGAGGCTGCTCCGCGCATGAAGTGCAGTCCAAAGCCGACA	1560
QY	1561	GCGGTGCATCTCGTGGTAAACCGCCGCTCTTTCTTCTCTCTCTCTCTCTCTCTCT	1620
DB	1561	GCGGTGCATCTCGTGGTAAACCGCCGCTCTTTCTTCTCTCTCTCTCTCTCTCTCT	1620
QY	1621	TCGGTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGG	1680
DB	1621	TCGGTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGG	1680
QY	1681	GAGATGGGAAAGATTGCTGTGGCTTGTGGGATGTTGCGTTAGATCCGTGCAAGCTG	1740

Db 1681 GAGATGGGAAAGATTCTGTGCTTGTGGGATGTTGCGTTAGATCGTGCAAGCTG 1740  
Qy 1741 TCTCCCTTTCTTTTGGGTCAAACTCGGTTTCTTCTTCGCTCTCTCGTGGTTTTT 1800  
Db 1741 TCTCCCTTTCTTTTGGGTCAAACTCGGTTTCTTCTTCGCTCTCTCGTGGTTTTT 1800  
Qy 1801 TGCAAAACAAAATTTTGGCTTCGAGGGAATCTTGTAATAATTTACTTGTGTTTCCACC 1860  
Db 1801 TGCAAAACAAAATTTTGGCTTCGAGGGAATCTTGTAATAATTTACTTGTGTTTCCACC 1860  
Qy 1861 TTTGGAACTGCGCAATTTTTCATGGTCAAAATTCGGTTTCTTCGCAACTTCGCAAT 1920  
Db 1861 TTTGGAACTGCGCAATTTTTCATGGTCAAAATTCGGTTTCTTCGCAACTTCGCAAT 1920  
Qy 1921 GAGTCGATTTGGGTTTTTCCATTTGACGGAAGGATATTAGTGACAGAATCAAGCGAAAAG 1980  
Db 1921 GAGTCGATTTGGGTTTTTCCATTTGACGGAAGGATATTAGTGACAGAATCAAGCGAAAAG 1980  
Qy 1981 AATATTTATATGATCGGTTCGACGCATTAATAGCAATTCAGGCGTTAAGGTATAGA 2040  
Db 1981 AATATTTATATGATCGGTTCGACGCATTAATAGCAATTCAGGCGTTAAGGTATAGA 2040  
Qy 2041 ATTATTTCTGATTTCACTAGTATGTTGAAATTTGAGGTTGCAATTTTGTCTCACACA 2100  
Db 2041 ATTATTTCTGATTTCACTAGTATGTTGAAATTTGAGGTTGCAATTTTGTCTCACACA 2100  
Qy 2101 TGTATTTTAAATTTGTCATGCGCAATGGTGTGAGGATGACGAAGAAACAAAGGCG 2160  
Db 2101 TGTATTTTAAATTTGTCATGCGCAATGGTGTGAGGATGACGAAGAAACAAAGGCG 2160  
Qy 2161 GGTTCGTGGTTTTTGTCTCATGTGAAACTTTTAAAAATAACCAAAACAGAAAGTGCA 2220  
Db 2161 GGTTCGTGGTTTTTGTCTCATGTGAAACTTTTAAAAATAACCAAAACAGAAAGTGCA 2220  
Qy 2221 ACTTTATATGATCGATAAAGGTTTGGAAATGTCACATCGAATAAATAAAGAGATAGA 2280  
Db 2221 ACTTTATATGATCGATAAAGGTTTGGAAATGTCACATCGAATAAATAAAGAGATAGA 2280  
Qy 2281 CGTATTACTCTTTTGTCTCAATATTCACACTAAACCAAAATCGATCTGATGTTTATGTA 2340  
Db 2281 CGTATTACTCTTTTGTCTCAATATTCACACTAAACCAAAATCGATCTGATGTTTATGTA 2340  
Qy 2341 AACTGTGTGACACAAATTTCAATTAATTAACCTTGTCTGTGGCAATGCTCTGCTTACG 2400  
Db 2341 AACTGTGTGACACAAATTTCAATTAATTAACCTTGTCTGTGGCAATGCTCTGCTTACG 2400  
Qy 2401 CAATGCACTACTCTAAATTTTACTGGATCATCAAAACAAATCATTTGGAAGGTTTATCTA 2460  
Db 2401 CAATGCACTACTCTAAATTTTACTGGATCATCAAAACAAATCATTTGGAAGGTTTATCTA 2460  
Qy 2461 TTGCTTCTTTTACGTTTCATGAAATAGCTAAATTTTCAAACTCTCTGCACTACCTAAATTT 2520  
Db 2461 TTGCTTCTTTTACGTTTCATGAAATAGCTAAATTTTCAAACTCTCTGCACTACCTAAATTT 2520  
Qy 2521 TAGTGGACCGAATCTAGTCATCGTTCGCGTTAGCTATCTCCAAATTTGATGGCGTTCT 2580  
Db 2521 TAGTGGACCGAATCTAGTCATCGTTCGCGTTAGCTATCTCCAAATTTGATGGCGTTCT 2580  
Qy 2581 AATTATGCAATGTCATCGAGAACTGGTGTGTTAAACAGCTTGAAGATGCGAGTAC 2640  
Db 2581 AATTATGCAATGTCATCGAGAACTGGTGTGTTAAACAGCTTGAAGATGCGAGTAC 2640  
Qy 2641 ACCCTAATATGTTATTCAGGATAGGATGACAGTGAATTAACATAGACAGAGTGTATGCAA 2700  
Db 2641 ACCCTAATATGTTATTCAGGATAGGATGACAGTGAATTAACATAGACAGAGTGTATGCAA 2700  
Qy 2701 TTGATGCTTTAGATGATTCATAGCATCTGCGAGAGCTTGAATTTATGATGTTAGATGG 2760  
Db 2701 TTGATGCTTTAGATGATTCATAGCATCTGCGAGAGCTTGAATTTATGATGTTAGATGG 2760  
Qy 2761 TAAGCAATCTCTGCGTTCTTAAAAAAGGTTAGACATCAATTTTGACATCCC 2820  
Db 2761 TAAGCAATCTCTGCGTTCTTAAAAAAGGTTAGACATCAATTTTGACATCCC 2820

Db 2761 TAAGCAATCTCTGCGTTCTTAAAAAAGGTTAGACATCAATTTTGACATCCC 2820  
Qy 2821 TTTATTTTACTAAATTTTAGTCCACCTGGCTGGGAAAGGAAACACAGTCCACGCTGAT 2880  
Db 2821 TTTATTTTACTAAATTTTAGTCCACCTGGCTGGGAAAGGAAACACAGTCCACGCTGAT 2880  
Qy 2881 TAAGATGAATTTTCTGCTGTCGCAATTTAGCCACTGGTGATATGTTGAGGGCTGAGTGGC 2940  
Db 2881 TAAGATGAATTTTCTGCTGTCGCAATTTAGCCACTGGTGATATGTTGAGGGCTGAGTGGC 2940  
Qy 2941 TGCTAAAACTCCACTTTGGGATTAAGGCTAAAGAAAGCTATGCAAGAGGTAGTTTAAAGAA 3000  
Db 2941 TGCTAAAACTCCACTTTGGGATTAAGGCTAAAGAAAGCTATGCAAGAGGTAGTTTAAAGAA 3000  
Qy 3001 ACATATAGCAACAGAAAATTAACACAGCAGGAATGGGTTTCTTGATTTCTTTTCTTTT 3060  
Db 3001 ACATATAGCAACAGAAAATTAACACAGCAGGAATGGGTTTCTTGATTTCTTTTCTTTT 3060  
Qy 3061 CCTTATCTTCTAGGGAGAGCTTGTTCCTGATGACTTGGTCTTGGGATTAATGATGAAGC 3120  
Db 3061 CCTTATCTTCTAGGGAGAGCTTGTTCCTGATGACTTGGTCTTGGGATTAATGATGAAGC 3120  
Qy 3121 CATGAAGAAAACTTCATGCCAGAAAAGGTTTATCCTTGATGGTTTCCCTAGAACTGTGT 3180  
Db 3121 CATGAAGAAAACTTCATGCCAGAAAAGGTTTATCCTTGATGGTTTCCCTAGAACTGTGT 3180  
Qy 3181 TCAAGCACAGAAAGGTGAGGTCCTTGCTCAATATGACCGCTATATAAAGAGCTCCCTTTT 3240  
Db 3181 TCAAGCACAGAAAGGTGAGGTCCTTGCTCAATATGACCGCTATATAAAGAGCTCCCTTTT 3240  
Qy 3241 TGTATTTAGAGCTGTCTATATAAATGGACAGTTTCTATCATTTGTATCATCTTTCTTACTA 3300  
Db 3241 TGTATTTAGAGCTGTCTATATAAATGGACAGTTTCTATCATTTGTATCATCTTTCTTACTA 3300  
Qy 3301 AAAAAATGGTGCAGCTTGTATGAAATGTCGCAACAAAGGTACTAAGATTGACAAAGTTCT 3360  
Db 3301 AAAAAATGGTGCAGCTTGTATGAAATGTCGCAACAAAGGTACTAAGATTGACAAAGTTCT 3360  
Qy 3361 AAAATTTGCAATTTGATGATGCAATACTCGAAAGAACAAATTCGCGTCTGATGCCACCC 3420  
Db 3361 AAAATTTGCAATTTGATGATGCAATACTCGAAAGAACAAATTCGCGTCTGATGCCACCC 3420  
Qy 3421 ATCAAGTGTGATGATCTATACATAAATTTGCTCCTCAAGACTCTCGACCTTGATGA 3480  
Db 3421 ATCAAGTGTGATGATCTATACATAAATTTGCTCCTCAAGACTCTCGACCTTGATGA 3480  
Qy 3481 TGTAAGTCAATACAGATTAATTTGCTCTCGTTGCAATTTGTGAGATCTCAGATTTTATAC 3540  
Db 3481 TGTAAGTCAATACAGATTAATTTGCTCTCGTTGCAATTTGTGAGATCTCAGATTTTATAC 3540  
Qy 3541 CATTTTTCATTTTCTAATAGATTTGGTACATATAAATGGTTGATGCTTGACAGATATGC 3600  
Db 3541 CATTTTTCATTTTCTAATAGATTTGGTACATATAAATGGTTGATGCTTGACAGATATGC 3600  
Qy 3601 GCCTTTTACCATAAATTTCCCTGCTCATCATCGATCATCGTAGTCTCTGCTTACTTGTGT 3660  
Db 3601 GCCTTTTACCATAAATTTCCCTGCTCATCATCGATCATCGTAGTCTCTGCTTACTTGTGT 3660  
Qy 3661 TTTTAAAGAACAAACATGAGCCATCTTCTTCAAAAATAAAGGTTTCAATTTGAGTACCT 3720  
Db 3661 TTTTAAAGAACAAACATGAGCCATCTTCTTCAAAAATAAAGGTTTCAATTTGAGTACCT 3720  
Qy 3721 TCCTCTTGGCAGCTGTTTAAATAGTTTGTCTTCCCAATAGTTTAAATAGGTAGTGATAT 3780  
Db 3721 TCCTCTTGGCAGCTGTTTAAATAGTTTGTCTTCCCAATAGTTTAAATAGGTAGTGATAT 3780  
Qy 3781 GGATATATTTTCTGTTTGGTGTCTCGTACGAAGAGTAAATGCACCTTTTGTG 3840  
Db 3781 GGATATATTTTCTGTTTGGTGTCTCGTACGAAGAGTAAATGCACCTTTTGTG 3840  
Qy 3841 ACAAGAAATGATAGGAGGTTTGTGATGACACATTTTGTCTTCTGCAATATG 3900  
Db 3841 ACAAGAAATGATAGGAGGTTTGTGATGACACATTTTGTCTTCTGCAATATG 3900

```
QY 3901 CCGTGGTAAAGTGGCCATATAAGATAGATACATGACATGTTTGTGGAGGAGATGTTGATA 3960
Db 3901 CCGTGGTAAAGTGGCCATATAAGATAGATACATGACATGTTTGTGGAGGAGATGTTGATA 3960
QY 3961 TTGTTTCTATTGCTTCCAAATTAAGAGATATCTCTTTAGCAATGATTTTCATGTTGGGAC 4020
Db 3961 TTGTTTCTATTGCTTCCAAATTAAGAGATATCTCTTTAGCAATGATTTTCATGTTGGGAC 4020
QY 4021 ATATTGTCGCTGCTATTAAGTAATTTGTTGATATCATATATATCTTTTAATTTGTTAAT 4080
Db 4021 ATATTGTCGCTGCTATTAAGTAATTTGTTGATATCATATATATCTTTTAATTTGTTAAT 4080
QY 4081 ATATTGTCGCTGCTATTAAGTAATTTGTTGATATCATATATATCTTTTAATTTGTTAAT 4140
Db 4081 ATATTGTCGCTGCTATTAAGTAATTTGTTGATATCATATATATCTTTTAATTTGTTAAT 4140
QY 4141 TTAATTCAGGAAGATGACACAGCTGCGATTTGAAGTCAAGGCTTGAAGCTTCCAC 4200
Db 4141 TTAATTCAGGAAGATGACACAGCTGCGATTTGAAGTCAAGGCTTGAAGCTTCCAC 4200
QY 4201 GTACAACTAAGCCCTGATGTTTCTTTAGCACTACGTTTTTAAATATTTCAGATATTCT 4260
Db 4201 GTACAACTAAGCCCTGATGTTTCTTTAGCACTACGTTTTTAAATATTTCAGATATTCT 4260
QY 4261 TTTAGGATGATGCTACTTCAGTTAAGGCGGATTCCTTCAGTTGCAATACAGTGTTC 4320
Db 4261 TTTAGGATGATGCTACTTCAGTTAAGGCGGATTCCTTCAGTTGCAATACAGTGTTC 4320
QY 4321 TGTATATCTTTCATGTTT 4341
Db 4321 TGTATATCTTTCATGTTT 4341

RESULT 2
CNS08C85 98782 bp DNA linear HTG 10-JUL-2002
LOCUS Oryza sativa chromosome 12 clone OJ1220_D01, *** SEQUENCING IN
DEFINITION PROGRESS ***, in ordered pieces.
ACCESSION AL731756
VERSION AL731756.3 GI:21732197
KEYWORDS HTG; HTGS_PHASE2; HTGS_ACTIVEFIN.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 98782)
Choise, N., Orjeda, G., Cattolico, L., Demange, N., Wincker, P.,
Seguren, B., Pelletier, E., Scarpelli, C., Sallanoubat, M.,
Weissenbach, J., and Quetier, F.
Oryza sativa chromosome 12 sequencing
Unpublished
2 (bases 1 to 98782)
Genoscope.
Direct Submission
Submitted (09-JUL-2002) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE [E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr]
On Jul 11, 2002 this sequence version replaced gi:20372820.
IMPORTANT: This sequence is unfinished and does not necessarily
represent the correct sequence.
Work on the sequence is in progress and the release of this data is
based on the understanding that the sequence may change as work
continue. The sequence may be contaminated with foreign sequence
from E.coli, yeast, vector, phage, etc.
The nucleotide sequence of this BAC clone was generated by
combining Monsanto and Genoscope sequencing data.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
Location/Qualifiers
```

```
source 1..98782
/organism="Oryza sativa"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:4530"
/chromosome="12"
/clone="OJ1220_D01"
/clone_lib="Monsanto"
BASE COUNT 30065 a 19192 c 20278 g 29238 t 9 others
ORIGIN
Query Match 91.6%; Score 3976.8; DB 2; Length 98782;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 4254; Conservative 0; Mismatches 72; Indels 44; Gaps 19;
QY 1 CTGCAGGAAGATTAATTAGGTGGACACACAAACCCCTGGTGGTGGACGCCCTGTTGT 60
Db 63789 CTGCAGGAAGATTAATTAGGTGGACACACAAACCCCTGGTGGTGGACGCCCTGTTGT 63848
QY 61 AATCAACTGGGTGTTGCTGGTGGACATGTTTTCAGGAAAAATTAAGCAAGAAAAATTAAG 120
Db 63849 AATCAACTGGGTGTTGCTGGTGGACATGTTTTCAGGAAAAATTAAGCAAGAAAAATTAAG 63908
QY 121 AAGATCTCTCAAGCTGACATGAGAAAAACGTAATCCAAATGGAAGCGAATTTCAAGTCGTT 180
Db 63909 AAGATCTCTCAAGCTGACATGAGAAAAACGTAATCCAAATGGAAGCGAATTTCAAGTCGTT 63968
QY 181 TCTTGTAATACCAATGTTTAGAATACATAAGACAGTGCACAC - GTTTGATGGCTCTTATTG 239
Db 63969 TCTTGTAATACCAATGTTTAGAATACATAAGACAGTGCACAC - GTTTGATGGCTCTTATTG 64028
QY 240 GCTCGTGTGATCTGACTTGTGTGACAAAGCATCAAAATGCTTCTTGGAGTATCTTTATT 299
Db 64029 GCTCGTGTGATCTGACTTGTGTGACAAAGCATCAAAATGCTTCTTGGAGTATCTTTATT 64088
QY 300 ACCGAAACCCCAAGATTAATCTATTCACCTCAGGGTAAATGTCGTAATATGCAAT 359
Db 64089 ACCGAAACCCCAAGATTAATCTATTCACCTCAGGGTAAATGTCGTAATATGCAAT 64148
QY 360 GAATACAAATTCGCAAAATATCATGTTTATCTATCTGTCATAATTTGAAATTTGAGTCCA 419
Db 64149 GAATACAAATTCGCAAAATATCATGTTTATCTATCTGTCATAATTTGAAATTTGAGTCCA 64208
QY 420 ACTGAGACTGCAATACGATTTTCTTTTCAAAAGAAATTAATAATTTTTCATGAA 479
Db 64209 ACTGAGACTGCAATACGATTTTCTTTTCAAAAGAAATTAATAATTTTTCATGAA 64268
QY 480 ACCCAATTCACCGTTCGAGAAATATGCTGTCATAAATAGTAGTCTAGTCGAGAAACAA 539
Db 64269 ACCCAATTCACCGTTCGAGAAATATGCTGTCATAAATAGTAGTCTAGTCGAGAAACAA 64328
QY 540 AATTAATATCACATAAAAAAGAGGTTGTTTAATTAATAAACCATTGTTTCGTACTACAATC 599
Db 64329 AATTAATATCACATAAAAAAGAGGTTGTTTAATTAATAAACCATTGTTTCGTACTACAATC 64388
QY 600 TAATTTGTAATTTCTTATTTTCAGTCACAAATTCCAATTTCCAAATTAAGAAAAATTAACG 659
Db 64389 TAATTTGTAATTTCTTATTTTCAGTCACAAATTTCCAAATTTCCAAATTAAGAAAAATTAACG 64448
QY 660 TAGACGGCTAAGCCCAACCCATCTAAGGCTAAGTTCGAGAGGTGAAGTACGCAAGAAAAAT 719
Db 64449 TAGACGGCTAAGCCCAACCCATCTAAGGCTAAGTTCGAGAGGTGAAGTACGCAAGAAAAAT 64508
QY 720 ATGATGTTTATTAATATGATTTTTTTAAATCACTTTTCACATAAATTTCTTTAGGAAC 779
Db 64509 ATGATGTTTATTAATATGATTTTTTTAAATCACTTTTCACATAAATTTCTTTAGGAAC 64568
QY 780 ATATCATTTAATGTTTGAAGAACCGTGCACATAAGAAAACTAAGAACGATGAGTTGGAA 839
Db 64569 ATATCATTTAATGTTTGAAGAACCGTGCACATA - GAAAACTAGAACGATGAGTTGGAA 64626
QY 840 ACAGAGAAAAACACAGCCTTAAGGCTTCTTGATCTCTAGTGGAGGTTGATTTTCAA 899
```

Db 64627 ACAAGAGAAAAACACAGCCTTAAAGGCTTCTTGATCCTCTAGTTGGAGGTTGATTTTCAAA 64686  
Qy 900 CGCATGATTAACGAGAAAGCTCAATTAGCACATATTACTTAGATATTTATTAATTTATTAAC 959  
Db 64687 CGCATGATTAACGAGAAAGCTCAATTAGCACATATTACTTAGATATTTATTAATTTATTAAC 64746  
Qy 960 TTGAAAAAATATTTATTTGAAATTTTAAACAATGTATGCAATAAATTTATTTTAAAAA 1019  
Db 64747 TTGAAAAAATATTTATTTGAAATTTTAAACAATGTATGCAATAAATTTATTTTAAAAA 64806  
Qy 1020 CACACCAATTTAACCTTTAAAAAGCATCTTAATAGGAAACGAGGAAGTTAAAGATTCAC 1079  
Db 64807 CACACCAATTTAACCTTTAAAAAGCATCTTAATAGGAAACGAGGAAGTTAAAGATTCAC 64866  
Qy 1080 CGAA--GTGTTGGATAAT-GAAAAATGGGGTGGGATTAGAAATTTGGTAAATGAATCAGG 1135  
Db 64867 CGAAGGTGTGTTGGATAATGGAAAAATGGGGTGGGATTTAGAAATTTGATGAATGAATCAGG 64926  
Qy 1136 GTTAGGATTAATATTAATAATGAAGAGGAGAAATGAATGGTTAGAGTTTAAATGTGTCT 1195  
Db 64927 GTTAGGATTAATATTAATAATGAAGAGGAGAAATGAATGGTTAGAGTTTAAATGTGTCT 64986  
Qy 1196 TTTTGGTGGGTAGAAAAATTTATTTGCCATACACTCCCGAGAGCGGTGCGTTCGCTGG 1255  
Db 64987 TTTTGGTGGGTAGAAAAATTTATTTGCCATACACTCCCGAGAGCGGTGCGTTCGCTGG 65046  
Qy 1256 GCAGAGCGGTCTTTTGGTGGAAAAAATACTGCTTAAAAAGGAAACAGAAAGGCCCA 1315  
Db 65047 GCAGAGCGGTCTTTTGGTGGAAAAAATACTGCTTAAAAAGGAAACAGAAAGGCCCA 65106  
Qy 1316 GCTTTT-GTTTGCACCGTCTCACAGAAACGAAACAAAAAGCCCAACCACTTAACCTCCT 1374  
Db 65107 GCTTTGGTTGTCACCGTCTCACAGAAACGAAACAAAAAGCCCAACCACTTAACCTCCT 65166  
Qy 1375 CGATCCGACGAGACTCCTCAATTCAGCGCGCACGCGGAGACACGCGAGCGGTTC 1434  
Db 65167 CGATCCGACGAGACTCCTCAATTCAGCGCGCACGCGGAGACACGCGAGCGGTTC 65226  
Qy 1435 GTCACGAAACAGGCTAGTGCAGTAGTTGTTGCTCGGAGATGGCGGCAACCTGGAGGA 1494  
Db 65227 GTCACGAAACAGGCTAGTGCAGTAGTTGTTGCTCGGAGATGGCGGCAACCTGGAGGA 65286  
Qy 1495 CGTCCGTCGATGAGCTGATGACGAGCTGCTCGCGCGCACGAGTGCAGCTCCAGCC 1554  
Db 65287 CGTCCGTCGATGAGCTGATGACGAGCTGCTCGCGCGCACGAGTGCAGCTCCAGCC 65346  
Qy 1555 CGACAAGCGGTCACTCCTCGTTAAGCCCGCCCTCTTCTTCTCCTCTCTCTCTC 1614  
Db 65347 CGACAAGCGGTCACTCCTCGTTAAGCCCGCCCTCTTCTTCTCCTCTCTCTCTC 65406  
Qy 1615 TCTCTCTCGGTGGTTGGTTTGGTTGGTTGGTGGATCCGTTCCGCGCGCGGTGCTGCCCT 1674  
Db 65407 TCTCTCTCGGTGGTTGGTTTGGTTGGTTGGTGGATCCGTTCCGCGCGCGGTGCTGCCCT 65466  
Qy 1675 GGTCCGAGATGGGGAAGATTCCTGTGCTTCTGGGATGTTGCGTTAGATCCGTGCA 1734  
Db 65467 GGTCCGAGATGGGGAAGATTCCTGTGCTTCTGGGATGTTGCGTTAGATCCGTGCA 65526  
Qy 1735 AAGCTGTCTCCTCTCTCTTTTGGGTGCAAACTCGGGTTCTCTCTCCGCTCCTCGTG 1794  
Db 65527 AAGCTGTCTCCTCTCTTTTGGGTGCAAACTCGGGTTCTCTCTCCGCTCCTCGTG 65586  
Qy 1795 GTTTTTTGCACAAATAATGTTGGCTTCGAGGAAATCTTGTAATAATTTACTTTGTTT 1854  
Db 65587 GTTTTTTGCACAAATAATGTTGGCTTCGAGGAAATCTTGTAATAATTTACTTTGTTT 65646  
Qy 1855 TCCACCTTTTGAACCTGCGCAATTTTTCATGCTGCAATTTTGGTTGCTTCTTCACTTT 1914  
Db 65647 TCCACCTTTTGAACCTGCGCAATTTTTCATGCTGCAATTTTGGTTGCTTCTTCACTTT 65705  
Qy 1915 CGGATTGAGTCGATGGGGTTTTTCATTTGACGGAAGATATTAGTGACAG-AATCAAG 1973  
Db 65706 CGGATTGAGTCGATGGGGTTTTTTCATTTGACGGAAGATATTAGTGACAGAAATCAAG 65765

Qy 1974 CGAAAGAAATATTTTATATGATCGGTGTTGACGCAATATAGGAATTTTACGGCGTTTAAG 2033  
Db 65766 CGAAAGAAATATTTTATATGATCGGTGTTGACGCAATATAGGAATTTTACGGCGTTTAAG 65825  
Qy 2034 GTATAGAAATTTATTCGTATATTCAGTAGTAGTGTGTAATTCGAGAGTTGCAATTTTGTTC 2093  
Db 65826 GTATAGAAATTTATTCGTATATTCAGTAGTAGTGTGTAATTCGAGAGTTGCAATTTTGTTC 65885  
Qy 2094 TCACACATGTTATTTTAAATTTTCAATGTCATGTCGAATGGTGTGAGATGACGCAAGAAACA 2153  
Db 65886 TCACACATGTTATTTTAAATTTTCAATGTCATGTCGAATGGTGTGAGATGAGC-AGAAACA 65942  
Qy 2154 ACAGGCGGTTCTGTGTTTTTGTCTGATGTGGAACCTTTTTTAAAAATAACCAAAACAGAA 2213  
Db 65943 ACAGGCGGTTCTGTGTTTTTGTCTGATGTGGAACCTTTTTTAAAAATAACCAAAACAGAA 66002  
Qy 2214 AGGTGCAACTTTATATGATATGATTAAGGTTTGGAAATGTCACCTCSGAATAAAAACAAAG 2273  
Db 66003 AGGTGCAACTTTATATGATATGATTAAGGTTTGGAAATGTCACCTCSGAATAAAAACAAAG 66062  
Qy 2274 AGATAGACGTTATTCCTCTTTTCAATATTCACCTACTAAACCAAAATCGATCTGTATGG 2333  
Db 66063 AGATAGACGTTATTCCTCTTTTCAATATTCACCTAGTAAACCAAAATCGATCTGTATGG 66122  
Qy 2334 TTATGTAACCTGTGCACACAAATTTTCAATATAAATCTTCTGTTGCGCAATGCTCT 2393  
Db 66123 TTATGTAACCTGTGCACACAAATTTTCAATATAAATCTTCTGTTGCGCAATGCTCT 66182  
Qy 2394 GCTTCAGCAATGCACTACTCTAAATTTTACTCGATCATCAAAACAAATCATTTATGGAAGTT 2453  
Db 66183 GCTTCAGCAATGCACTACTCTAA--TTTACTGGTCAATCAAAACAAATCATTTATGGAAGTT 66240  
Qy 2454 TATTTCTATTCGTTCTTTTACGTTTCATGAAGTATAGCTAAATTTTACAACTCTCTGCACATC 2513  
Db 66241 TATTTCTATTT--GTTCTTTACTTCATGAAGTATAGCTAAATTTTACAACTCTCTGCACATC 66298  
Qy 2514 CTAATTTTATGAGACGAAATCTAGTCATCGTCTGCGTTAGCTATCTCCAAATTTGATGGC 2573  
Db 66299 CTAATTTAAGTGACGCAATCTAGTCATCGTCTGCTCTTAGCTATCTCCAAATTTGATGGC 66358  
Qy 2574 GTGTCTCTAATTCATGATCATGTCACCTGAGA-ACTGGTGTGTTAACAGCATTTGTAAGAAT 2632  
Db 66359 TTGTTCTAATTCATGATCATGTCACCTGAGA-ACTGGTGTGTTAACAGCATTTGTAAGAAT 66418  
Qy 2633 GCCAG-TACACCTAATATGTTATTTGAGATAGGATAGTG-----A 2672  
Db 66419 GCCAGTTACACCTAATATGTTATTTGAGGATAGGATAGTTGACAACTTGCATATCCATA 66478  
Qy 2673 CAGTGTAAACATAGAGCAAGTGTATGCAATTTGATGCTTTTAGATAGTTTCATAGCATCTGGCA 2732  
Db 66479 GAGCAAAAGTGATGTCAAATTTTGATGCTTTTATGGAATGTTTAAATAGGCATCTGGCAG 66538  
Qy 2733 GAAGCTTGAATATTTGATG-TCTAGAAATGGTAAGACAACTCTCTGCGTGTCTTAAAAAAAAA 2791  
Db 66539 AAAGCTTGAATATTTGATGTTCTAGAAATGGTAAGACAACTCTCTGCGTGTCTTAAAAAAAAA 66598  
Qy 2792 AAAGTTAAGACATACAAATTTTCACATCCCTTTATTTACTAAATTTTAGGTCCACCTGGC 2851  
Db 66599 AAAGTTAAGACATACAAATTTTGAATCCCTTTATTTACTAAATTTTAGGTCCACCTGGC 66658  
Qy 2852 TGCAGAAAGGAAACACAGTCACCGCTGATTAAGGATGAATTTTCTGTTGCGCAATTTAGCC 2911  
Db 66659 TGCAGAAAGGAAACACAGTCACCGCTGATTAAGGATGAATTTTCTGTTGCGCAATTTAGCC 66718  
Qy 2912 ACTGGTGATATGTTGAGGCTGACGTGCTGCTAAAACCTCCACTTTGGGATTTAAGGCTTAA 2971  
Db 66719 ACTGGTGATATGTTGAGGCTGACGTGCTGCTAAAACCTCCACTTTGGGATTTAAGGCTTAA 66778  
Qy 2972 GAAGCTATGGAAGAAGTATGTTTTTAAAGAAACATATATAGCAACAGAAATTTATAACAGCAGG 3031  
Db 66779 GAAGCTATGGAAGAAGTATGTTTTTAAAGAAACATATATAGCAACAGAAATTTATAACAGCAGG 66838





```
/product="4-coumarate-CoA ligase-like protein"
/notes="CDS is reported in Acc# AB023035
gene id:K9H21.8"
/number=2
/evidence=not experimental
complement(625..1680)
/product="4-coumarate-CoA ligase-like protein"
/notes="CDS is reported in Acc# AB023035
gene id:K9H21.8"
/number=1
/evidence=not experimental
complement(join(2166..2408,2476..2734,2838..3121,
3189..3278,3360..3528,3611..4245))
/notes="gene id:MLE2.2"
/codon_start=1
/evidence=not experimental
/product="auxin-independent growth promoter-like protein"
/protein_id="BAB08804.1"
/db_xref="GI:9758280"
/translation="WAKSMRTIKNPFTTSHPPSPFHLILLFTPLKSSPRVNLHHH
QRCPPQFLLISLVFSGISFLTSLDTGSTCVSTSSASLKFIFISDVDRQI
LTSVAGASTLUPFAPAGGHGNNTEEEKFQKOPGEGYKCLDFSLSYKKSASV
SEKRFVLVVVGLNNOQNIQVAVIAMLEALVVPVQVQVNRVNGDESEFSLF
DVEHKTLKSDVRVTSLSPLTHLMSRQTIENQIPWQSPVWIRAKYKQVLEGLIV
EKGLSKLAKNPDLQKLRCKVAFHARFAAPENLGNKLTIRRMTEGYIAHLRL
EKQVWRTGCLGLSEPDRIIAETRTSQPRLTGRNLTYTERRLAGFCPLNVETIA
RLKALGAPNSAIYIAGEPEPGSGRALEPLAKFSNLVTKETLAHKGELLPTNRS
ALAAIDYVLSDDVFISSHGGNMAKQGNRAVYGRKFPMPNKRAMLPMENSSVS
DAELFPLTKLRKSGQGPESRRGRDRDVIAYVPECMCRKRRKRSVGF"
complement(join(4928..5026,5118..5207,5299..5466,
5572..5691,5787..5932,6354..6471))
/notes="gene id:MLE2.3"
/codon_start=1
/evidence=not experimental
/product="adenylate kinase"
/protein_id="BAB08805.1"
/db_xref="GI:9758281"
/translation="WATGSGAADLEDVQTVLMSBELRLRKCQKPKRILFIPGPGS
GKGTQSPVVKDYCLHSTDMRAAVASKTPLGVKAKAMEKGLVSDLVVGIID
EAWNPCKOKGIFLQFPRTVTAQKLDMLKRGTEIDKVLNFAIDDAILERITGR
WIHPSSGRSYHTKFAPPKTPGVDDITGBPLQKRDADNLKSLAAPHQQTQPVIDY
YAKKAVLTNIOAEKAPQEVTSVKALS"
complement(join(6918..7036,7117..7245,7475..7648,
7793..7993,8083..8219,8312..9422))
/notes="contains similarity to receptor-like protein
kinase
gene id:MLE2.4"
/codon_start=1
/evidence=not experimental
/product="adenylate kinase"
/protein_id="BAB08806.1"
/db_xref="GI:9758282"
/translation="WGLFKFIFLLSLLMSFVSLSSQLQASQAVLLQLKHHLEYPQO
LESWYDHRNFYQLQATPSMNTFCFNSVSELTNIPGDSSEKAKSPFGF4ISNVLSD
GSIIESFVTLRLKSLRVLTLASIGWIRLPEKHLRSLSELYLDLNNFLFGSVPPK
LSTMVLETRFDHPNFNGTLPSWFDIVKLSFKSNLKGELSHLSLLSLSTIEYI
DIRANSLGSLPDLKCSKLFWDIDSNKLTGKLPRCLSSKQDIALRFNCNCISLEK
QHPSPFCYKVRAAKAEKAEAEANESGKRWKKKCALIGLIVGSMVLVLCVY
FILLRKGTGKTHVHNTVQDNPHTTGFSSSELNSNARTISRTSFGSEDLFVCRQSL
BEIVKATNFDMILGESSLYGTLYKGNLENGTKVAIRCLPSSKKYSIRNLKRLDL
LAKLRHNLVLCILGICDGGKDDYSVEKFLIYEYIENGFOQLSDNSSGKMMMS
ERNVLTVGAKAVHFLHTGVI PGFFSNRLKTNVLLNQHRFAKLSDYCNKRIANVKA
RRRLQMLWDSITNRTWITICIRKRSISSTRRTGIVGERRRKEGSDSTSHLSKRT
DKSDNSDEOMCT"
join(12169..13370,12719..12950,13056..13179,13364..13420,
13649..13840,13919..13996,14120..14302,14505..14585,
14791..14856,14952..15044,15143..15268,15478..15630)
/notes="gene id:MLE2.5
similar to unknown protein
sp|P54122"
/codon_start=1
/evidence=not experimental
/protein_id="BAB08807.1"
/db_xref="GI:9758283"

/translation="MMKPASLQGFSSHASSIYSDVRRPATTFSKMAAFSALSCLPYT
FTFRSSRIKSTVSCVTSASPSTSSSKTPRRSRLEGVSKMEDSVKRWEOFY
EGTDPPRLIPIGGLGEIWMNCLMVGNYDRYILIDAGIMPDPYDERGIQKIMPDTGF
IRRWKHKEAVITHGEDIHIGALPWIIPALDNPNTPIFASPTMELIKRLKEHGIPIV
OSRLKTESTRERPMAGPELEPIVTHSIIPDCSGLFLRCADGNILHVTGDKWIKDEAPLD
QKVFDRALBELSKEGVTLMNSDNTVLSPGRTISEKVDALPVNRVMAKGRVITQ
FASNIHRLGTSKAAADITGRKLVFGVMSLRTYLEAARDGKAPIDPSLSIKVEDIEAY
APKDLLVITTSQAEPPAANLASYSGHAFKTKEDILYSKAVIPGNSRSRWMMN
RIADIGNIIMGKNEMLHTSHAYRGELVLKIVKPQHFLPIHGELFLFKHEHLGKST
GIRHTTVLSLSEKTYELMVDK"
join(16291..16714,16783..17258,17369..17461)
/notes="gene id:MLE2.6
pir||T01257
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB08808.1"
/db_xref="GI:9758284"
/translation="MEIMRPGVSENTLKGRITITROMWLDKGRLLDALHKAHAALS
SCPVTCLSHMERTVSEVLKIVKYSKSPETIAIATENPMARADEVARLSGDPS
VGSVAALRKVVEGNDKSRKAKAPQEAQPKVEDRTLEDDIIDSARLLAEETAST
YTEEDTTPVSGSSSEDDFWKSFINPSSPSETENNNKVAOTPEPAKREKRSRDL
ELADASSETKSSPKRVKNNKPEEIKKIVRMGELHSRPQVVQGRMALWBEISSRL
SAEGNRSPOGCKSLWASLIQYEEKADERSKTSWPHFEDMNNILSELGTPAS"
join(18617..18789,18882..19035,19112..19208,19304..19392,
19521..19589,19790..19849)
/notes="gene id:MLE2.7
unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB08809.1"
/db_xref="GI:9758285"
/translation="MQSGILVCLYTNLDPLQDTLOQKMPKRTKDRSNVLDKKTHLA
RLNVSEGGKVLKRGEGMERQFMCNIGCELFVCYRAEENLETASFIYVDGALSVA
AAETNPQDAPVPCISQLDGLVOVAJEVEDRAORSATRVNADDVDTVAAPARGE
ANNELEFMWGRVLGRLSQMTLQGRKSKLLVVEDLSARQYVEKLEAVP"
complement(20641..22173)
/notes="gene id:MLE2.8"
/codon_start=1
/evidence=not experimental
/product="cytochrome P450-like protein"
/protein_id="BAB08810.1"
/db_xref="GI:9758286"
/translation="MEMLNAILLPIIGFVLIFSPPTTLKAKTASPSNPTSYQLI
GSILSFKNRHLRLLQWYTDLLRSPQTIITVDLLFGRRTIITANPENVEHLKTNFYN
PPIKGPFTDLLGLGGIIFNSGELMSQKSLASHEFTWRSLEFTEILREEVQR
LIPVLSAADCGETVDFQVLRFAFDVCKVSLGNDPDCDLTRPVPELVKAFDVA
BISARRATEPVAVVAVKVRVNSVSEKRLREAIKTVHLSVSEIIRAKKSLDIDGVS
DKQDLLSRFLAAGHEARVDSVIFSMAGRDTTSAAMTWLFWLLSQNDVDVETKILDE
LRNKSGLGFEDLREMSYTKACICEAMRLYPVAVWDSKHAANDDILPGCTPLKKGDK
VTIYPYGRMRMEKVGKWDDEFKRNRFEEBPSVGTVPULKSVSSEFPVFOGPRVC
IGKEMAFQMKYVVGVSVLRSFKIIPVCNNRPVFPVPLTAHMAAGLKVKIKRRQCDSM
YI"
complement(join(25307..25416,25470..25577,25667..25731,
25857..25980,26228..26273,26940..27053))
/notes="gene id:MLE2.9
unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB08811.1"
/db_xref="GI:9758287"
/translation="MEASSSQPSDSSQTSKFLSLPGRGFLSSSTVSSNPGSLRVY
ICEHPTSPQGLIKTQNTLIRSLLLKQKGSSEKSKDGKTGADCEPKKRAANRAJD
DRSSAKRAANASROGSSSRPTGERDFQSLVEKRLAMEKGLPTKGRKATSQSTLN
SILKMLSGVASSLSLFIVALVPLIDYGT"
complement(27986..28738)
/notes="gene id:MLE2.10"
/codon_start=1
/evidence=not experimental
/product="transcription factor Hap5a-like protein"
/protein_id="BAB08812.1"
/db_xref="GI:9758288"
```

```
Query Match      6.4%; Score 277.8; DB 8; Length 58527;
Best Local Similarity 65.4%; Pred. No. 9e-46;
Matches 464; Conservative 0; Mismatches 222; Indels 23; Gaps 3;

QY 2792 AAAGGTAAGACATACAACTTTTGACATCCCTTTATTTTACTAAATTTTAGTCCACCTGGC 2851
DB 5981 AATCGTTGGATGATGTTGATATGCTTTGTTTGGTGAAATTCGAGGACCTCCAGGG 5922

QY 2852 TGGCGAAGGAAACACAGTCACCGCTGATTAAGATGAATTTTGTGTCCTATTTAGCC 2911
DB 5921 TCAGGGAAGAGTACTCAATCTCCAGTAGTAGTGAAGGATGATTTGCTTGTGTCACTTATCC 5862

QY 2912 ACTGGTGATATGTTGAGGGCTGCAGTGGCTGCTAAACTCCACTTGGGATTAAGGCTAA 2971
DB 5861 ACTGAGACATGTTAAGAGCTGCTGTTCTTAAGACCCCTCTTGGTGTCAAGGCTAA 5802

QY 2972 GAAGCTATGGAACAGGTTAGTTTTTAAGAAACATATAGCAACAGAAATTAATACCAGCAG 3031
DB 5801 GAAGCTATGGAAGAAAGTCAGTCTCTTTTGTCTTTTGGAGAGATTTCTATTATTGTT 5742

QY 3032 AATGGGTTTCTTGATCTTTTGTGTTCTT-----TCCTTATCTTCTAGGAGAGCTTG 3083
DB 5741 CCTTTCCTGTTATAGATTTTATTTCTTTGTGATGTTGTTTTTATTTTCAGGGAGAGCTCG 5682

QY 3084 TTTCTGATGACTTGGTTGTTGGGATTTATGATGAAGCCATGAAGAAACTTTCATGCCAGA 3143
DB 5681 TCTCTGATGATTTGGTTGTTGGTATTAATGATGAAGCCATGAAGCCAAATGTCAA 5622

QY 3144 AGGTTTTATCTTGATGGTTTCCCTAGAACTGTTGTTTCAAGCACAGAAAGGTGAGGTCT 3203
DB 5621 AAGGATTTATCTTGATGGTTTCCAGGAGCTGTTACTCAGGCAGAGAGGTGATGCTCA 5562

QY 3204 TGGTCAATATGCACCGCTATATAAAGAGCTCCTTTTGTATTAGAGCTGCTATATA 3263
DB 5561 ATAGAAAACACTACTGTTGATC---GGTCTCTTTTGGTTCTTTGGTTTGAATGTAA 5505

QY 3264 ATGACAGTTTCTATCATGTTATCATCTTTTCTTACTAAAAAATGGTGCAGCTTGATGAA 3323
DB 5504 TTAATCTTCAACCATGTGCTGTCTTTGTT-----TGCAGCTCGATGAGA 5457

QY 3324 TGTGGCCAAACAGGTPACTTAAGATTGACAAGGTTCTTAATTTGCAATTGATGATGCA 3383
DB 5456 TGCTTAAGAGCGAGGAACTTGAATTTGACAAGTTCTTCAACTTTGCTATTGATGACGCA 5397

QY 3384 TACTGGAAGAACGAATTAACCGGTCGTTGGATCCACCCATCAAGTGGTAGATCTTATCATA 3443
DB 5396 TCTTGGAGGAAGAATAAACCGGGCGATGGATCCACCATCGAGTGCAGGAGTTACCACA 5337

QY 3444 CAAATTTGCTCTCTTACAGACTCCTGGACTTGATGATGATGATGATGATGATGATGAT 3492
DB 5336 CCAATTTGCTCTCTCCAAACCCCTGGAGTTGATGATGATGATGATGATGATGATGAT 5288

RESULT 4
RICADKA
LOCUS      Ricadka
DEFINITION Oryza sativa mRNA for adenylylate kinase-a, complete cds.
ACCESSION D10334
VERSION    D10334.1
KEYWORDS   adenylylate kinase-a; Adk-a.
SOURCE     Oryza sativa (strain:Yamahoushi) callus cDNA to mRNA,
           clone lib:lambda gt10 clone:Adk-a-29 and Adk-a-50.
ORGANISM   Oryza sativa
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
           Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE  Kawai,M., Kidou,S., Kato,A. and Uchimiya,H.
           Molecular characterization of cDNA encoding for adenylylate kinase of
           rice (Oryza sativa L.)
AUTHORS    Plant J. 2 (6), 845-854 (1992)
JOURNAL    93258424
MEDLINE    93258424
REFERENCE  2 (bases 1 to 1154)

Kawai,M.
Direct Submission
Submitted (30-JAN-1992) Maki Kawai, Univ. of Tokyo, Inst. of
Molecular and Cellular Biosciences, Yoyoi 1-1-1, Bunkyo-ku, Tokyo
113, Japan (Tel:03-3812-2111(ex.7845), Fax:03-3812-2910)
FEATURES
source
1..1154
/organism="Oryza sativa"
/strain="Yamahoushi"
/db_xref="taxon:4530"
/clone="Adk-a-29 and Adk-a-50"
/clone_lib="lambda gt10"
/dev_stage="callus"
1..1154
/gene="Adk-a"
137..862
/gene="Adk-a"
/EC_number="2.7.4.3"
/codon_start=1
/product="adenylylate kinase-a"
/protein_id="BAA01180.1"
/db_xref="GI:391877"
/translation="NAANLEDVPSMELMTELLRRMKSSKPKRVLVPGPCGKGKGTQ
SPLIKDFCLCHLATVDMRAAAKATPLGKIDKAKMDKBLVSDLVVGIIDAMKK
TSCKGFLDGPRTVVOAKLDEMLAQGTIKDKVLNFAIDDALEERITGRWTHPS
SGRSYHTKFAPKTPEGLDVTGEPLIQKDDTAALVKSRLEAFHVQTKPVIDYTKG
IVANLHAEKPKPEVTEVQKALS"
1078..1086
polyA_signal
1078..1086
/gene="Adk-a"
BASE COUNT 327 a 254 c 270 g 303 t
ORIGIN
Query Match      5.3%; Score 229; DB 8; Length 1154;
Best Local Similarity 99.2%; Pred. No. 5.1e-36;
Matches 241; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1337 CAGAAACGAAACAAAAAGCCCAACACCTCTCGATCCGACCGAGACTCCTCCA 1396
DB 1 CAGAAACGAAACAAAAAGCCCAACACCTCTCGATCCGACCGAGACTCCTCCA 60

QY 1397 TTTTCAGCGCGGACGGGAGAGACCGGAGGAGTTCCTCGACGACAGCTAGTGA 1456
DB 61 TTTTCAGCGCGGACGGGAGAGACCGGAGGAGCAGCGA--CGAGTTCGTCGACGACAGGCTAGTGA 118

QY 1457 GTAGTTGTTCTCTCGGAGATGGCGGAAACCTGGAGGAGCTGCCGTGATGGAGCTGATG 1516
DB 119 GTAGTTGTTCTCGGAGATGGCGGAAACCTGGAGGAGCTGCCGTGATGGAGCTGATG 178

QY 1517 ACGGAGCTGCTCGCGCGATGAAGTGCAGCTCAAGCCCGACAGGCGGTCTATCCTCGTC 1576
DB 179 ACGGAGCTGCTCGCGCGATGAAGTGCAGCTCAAGCCCGACAGGCGGTCTATCCTCGTC 238

QY 1577 GGT 1579
DB 239 GGT 241

RESULT 5
AB012248/c
LOCUS      Arabidopsis thaliana genomic DNA, chromosome 5, Pl clone:MXI22.
DEFINITION Arabidopsis thaliana
ACCESSION AB012248 BA000015
VERSION    AB012248.1 GI:3046856
KEYWORDS
SOURCE     Arabidopsis thaliana (strain:Columbia) DNA, clone lib:Mitsui P1
           clone:MXI22.
ORGANISM   Arabidopsis thaliana
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE  1 (sites)
AUTHORS    Kotani,H., Nakamura,Y., Sato,S., Asamizu,E., Kaneko,T., Miyajima,N.
           and Tabata,S.
```



**TITLE** Structural analysis of Arabidopsis thaliana chromosome 5. VI.  
**JOURNAL** Sequence features of the regions of 1,367,185 bp covered by 19  
**MEDLINE** physically assigned P1 and TAC clones  
**REFERENCE** DNA Res. 5 (3), 203-216 (1998)  
**AUTHORS** 98403884  
**TITLE** 2 (bases 1 to 80115)  
**JOURNAL** Nakamura, Y.  
 Submitted (23-Mar-1998) Yasukazu Nakamura, Kazusa DNA Research  
 Institute, Department of Plant Gene Research; 1332-3, Yana,  
 Kisarazu, Chiba 252-0812, Japan (E-mail: ynakamura@kazusa.or.jp,  
 Tel:81-438-52-3935, Fax:81-438-52-3934)  
**COMMENT** Address for correspondence: kaos@kazusa.or.jp  
 For the latest information on annotation of this clone, please see  
 http://www.kazusa.or.jp/kaos/cgi-bin/agd/graph.cgi?c=MX122  
 Genes with similarity to proteins in the databases are described in  
 'product' or 'note' qualifiers. Genes that have no significant  
 protein similarity are described as 'unknown protein'.  
 The software programs used to predict genes include: Grail  
 (Informatics Group, Oak Ridge National Laboratory,  
 http://compbio.ornl.gov/Grail-1.3/),  
 GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),  
 NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of  
 Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and  
 SplicePredictor (Volker Brendel, Stanford University,  
 http://grenlin.zool.iastate.edu/cgi-bin/sp.cgi).  
 Genes encoding tRNAs are predicted by tRNAscan-SE  
 (Sean Eddy, Washington University School of Medicine, St. Louis,  
 http://genome.wustl.edu/eddy/tRNAscan-SE/).  
 This sequence may not be the entire insert of this clone. It may be  
 shorter because we remove overlaps between neighboring submissions.  
 The 5' clone is K6A12 and the 3' clone is M6A10.

**FEATURES** Location/Qualifiers  
 source 1..80115  
 /organism="Arabidopsis thaliana"  
 /strain="Columbia"  
 /db\_xref="taxon:3702"  
 /chromosome="5"  
 /clone="MX122"  
 /clone\_lib="Mitsui Pi"  
 join(5147..5278,5501..5587,5690..5758,5863..5995,  
 6080..6126,6221..6307,6758..6819,7113..7155,7255..7335,  
 7421..7534,7731..7836,7918..7981,8527..8596,8687..8848,  
 8944..9072,9163..9312,9408..9494,9597..9665,9768..9866)  
 /note="Contains similarity to unknown protein  
 gb|AA949898.1  
 gene\_id:MX122.1"  
 /codon\_start=1  
 /evidence=not experimental  
 /protein\_id="BAB09450.1"  
 /db\_xref="GI:9758913"  
 /translation="MGKTKKPGKKEKTKAKADEKAKREGKLSPEDDIDAIL  
 LMQEAKKEVHVVEENVAAPSPRNSCSLTINPLKETELILYGFYNGQKTYVGD  
 LRYDVEQEWKLVSPSPRSHQAQVAKNYLIFGGEFTSPNQERPHYKDFWM  
 LDVKNQEWLNKGCPSRSHRMVLYKHKLIIFGGFYDLREVRVNDLYVFDLDQ  
 YKQEIKPQGMWPTARSQFFVQDEIFLYGYSKEVSEKSEKGIHVADLMSL  
 DRPTWENKVKIIGMPPSRAGSFVCHKRALIFGVDMEMEGDMVMSLFLNELYG  
 FOLDNRVSDKSKALQAKTRLDQVSVSEEVKPCGRINCMVVGKOTLYTYGGMME  
 IKDKVTLDLYLNLKSLDEWKCIIPPTTEWVEVSDDEGDDDDSDDEGENSE  
 ESDDDDDEEPMVDGSKVGVVAMIKGQKSLRRKRAIRIQIRANLGLSDSO  
 RPYRGEETKPKYKTNMYQMAAYEHTQHGKRLKQDFDLAEITYLELKLIDELA  
 ILBAQKAEAEAGSPASSRKGKGIKKR"  
 10440..11115  
 /note="contains similarity to mutator-like transposase  
 gene\_id:MX122.2"  
 /pseudo  
 /codon\_start=1  
 /evidence=not experimental  
 complement(join(11785..11943,12033..12185,12697..12801,  
 12893..13121,13152..13358,13545..13719,14023..14195,  
 14294..14363,14921..15378))  
 /note="gene\_id:MX122.3"  
 /codon\_start=1

/evidence=not experimental  
 /product="histone acetyltransferase"  
 /protein\_id="BAB09451.1"  
 /db\_xref="GI:9758914"  
 /translation="MNGELKKQPRGKGQYQGRGLTEEARVRAISEIVSMIRSRHR  
 NENVDLNAIKTACRKYGLARAPKLVEIMTAALPDSERETLLPKRAKVRANPVQVA  
 AVMSKPHRCPIATIGNICVYCPGPDSDFEVSTQSYGTEPTSRARVRYNVPVQA  
 RSRIDQKRLGHSDVKEFVILMGFTMSLPAEVRDFFIRNLHDALSGHTSANVEBAV  
 YSESAKICGMTIECPDYCLGPHLRQMLYVCTLEIGVOSTVEDADGKLYETLV  
 AAVADCCLAQDAGFKVNAHMPDLPNVGVVERMESFEFSPSPFRADGKLYETLV  
 IARGVELMKTGRINYPEQLVDIVARILSMVPEWTVRYVRQDIPMLVTSVEK  
 GNRLRLAARDMDLGLKCRDVTREAGIQDIHHKIKPQVELRVLRVAVGTAVHGRDADKLOHQ  
 YGTLMREARIRAREHRSNKIGVISGVGTRHYRYKGLVEGTPVPMVYKHL"  
 complement(join(16097..16288,16390..16543,17220..17342,  
 17531..17633,17815..17939,18018..18146,18527..18721,  
 18809..18937,19035..19148))  
 /note="contains similarity to ABC transporter  
 gene\_id:MX122.4"  
 /codon\_start=1  
 /evidence=not experimental  
 /protein\_id="BAB09452.1"  
 /db\_xref="GI:9758915"  
 /translation="MHGFDFKEIQEKISDSFRPQWRSFQVVRATNIYTVGYFOLRV  
 SLVDKAKQEMERQEOAADKIYFMCSDLGQFLKIAQLLAKPDMAAPAAWVKLVV  
 LCDQAPTPPDALIVLEKELSGICEIETPDEKPLGSAIAOVQHIGLIERLMTDI  
 ENLOALFYQRTDIKEDLHSITKEMKQIGVEYFDFKREANAMERIRCFLENNKSP  
 VIVPRVLDWTKVLYWETIYNGIPILSIGDEMAKKGINPHGKIAEAAKHNLNLSR  
 AYGMILKSGFFHADPHDPKILICKQEAHLEMLHTVAKCENEQOQLKLAOTLF  
 DTQMTQOTVLQPFSDSSIKKIAVTFPEELFVLTVTYVLRGLSGVNGVNSCAEO  
 WRANERALLASGRVTRGKWLRYVLTTLFA"  
 complement(join(22220..22387,22459..22531,22650..22732,  
 22903..22941,23062..23139,23402..23557,23788..23853,  
 24049..24129,24218..24280,24552..24659,24908..24991,  
 25355..25405,25557..25615,25689..25744,25838..26127))  
 /note="gene\_id:MX122.5"  
 /codon\_start=1  
 /evidence=not experimental  
 /product="DNA repair protein-like"  
 /protein\_id="BAB09453.1"  
 /db\_xref="GI:9758916"  
 /translation="MNGSKKGTVMVVCESQHSQWMSGRACHKVTGTMKRPSEG  
 SSSASGGNGTGLGTEGKGLSWLPEQATVQPHRLTDVIGHITQQQWRISLPLGFG  
 NEVARVGGGLAPGSLIIGDPCIGKSLILQGSLELAEPAPVLYSGEESVEQIGSR  
 ADNRVQTEELYLFSSDLQDILINKAHRSLSPRALIIDSITQVYLDKDVSGAGLTQVK  
 ECTSTLLRAKKSNIPLVGVHTKAGIAGPVLEHIVDVVLYLEGEBSYRLGRS  
 VNRFGSTDELGVFEMANSLEVSNPSGIYLSQQNPDSDLAGLAVAVMDGSKSL  
 IEVQALCPGSTVSRHVYVQASRADMIITAVLMKQAGRIQENGFILYANGMALSET  
 AGDLIAAAICSSFLFPIPHGVAFIGEIGLGEVTVPRMEKRVSTVAKLGFNKCIV  
 PKSVESLALSLKEIIGCKNLKELINAVPKG"  
 complement(join(29282..29509,29616..29945,30038..30175,  
 30265..30384,30548..30746,30938..31105,31525..32196))  
 /note="gene\_id:MX122.6"  
 /unknown protein  
 /codon\_start=1  
 /evidence=not experimental  
 /protein\_id="BAB09454.1"  
 /db\_xref="GI:9758917"  
 /translation="MATSADNSTMKRTTSLAKSESQSDSSYSTRSARRHRSLSRF  
 SHRPDPDIEVTVMRKGVNTVRGSGFGISLDDIAVEFPSPFSGESISGSGRGRS  
 CLRWSGTGVTEVTNSQRCRSVRVSGSGGGLRLDADTSSRRRSLSPASIN  
 NRGSSKVDVSVNNSSRRRSVRQPNRKYVENSQGVIEGDRENANSSRRSLSVRR  
 RIENSSDVDDQVQSSSRSDVKFMSGKQNSQKSAASDNKQGLRSFSQNPYIKH  
 DYGSSQSAVDDDEGSDSSSKGHTIRITVYAQNKAIPKKRSGNSGNSQKRLS  
 DNRGVTSTKGYATKLQESERKRELLAIMEQRGRELVLNKEILTENSBAE  
 KPLRTRKSKDRSRMSMCLTDEAEQIFIDEFTSNIEDTDFSLSEBSSESSFGMIK  
 TQSSQTSALKSTSVEMDGVMLQWETPDVSAVAACILKSPHTPNKSLVWESDIT  
 QDASGRGTSIGTSRSGSSPYDSVTKPWPTTKLKIDVAEYLKPNSSDILNETWK  
 LRHTSSGSLVLCSPSLI"  
 complement(36018..36893)  
 /note="emb|CAB62340.1  
 gene\_id:MX122.7  
 similar to unknown protein"

```
/codon_start=1
/evidence=not experimental
/protein_id="BAB09455.1"
/db_xref="GI:9758918"
/translation="MNFPGGCCIARYGGGGNDMSKVDRIMLRVRIAPRPDSGGSP
ASPTKSGVITNPSSRRKGRKYSKENNSSSTGSVNGNSKQRNDETKNGSGGG
REIYVLPPLPPTPKDQSLPKAKAALPGLAALWLSFNDGASYNRYOTELMTETWS
SLLVCEVTERLMEVEYLGCTDEERKMNLERDTCPGFISDGLGRVITNGSYRLVV
GKDHQCSKMSVLVMEKEPLVIRYRFTCRMRLQYTCRDKEYSSITSFCDVWRMSDGG
FAWLDVDAALCLGR"
complement(join(39967..40068,40144..40233,40331..40498,
40586..40851,41096..41216))
/note="gene_id:MX122.8"
/codon_start=1
/evidence=not experimental
/product="adenylate kinase"
/protein_id="BAB09456.1"
/db_xref="GI:9758919"
/translation="MATSSAASVDMEDIQTVLMSLELLRMKCSKPKRLVFIQPPG
SGKGTQSVINDEFCLHSLTGBMLRAAANKTPLGKAKEMDKGELVSDLLVVGIM
DEAWNPKQKGFILDGFPRTVTOAEKLDEMLNRGAQIDKVLNFAIDDSVLEERITG
RWIHPSSGRSYHTKFAEPKGVDDLTGELIQKODNADVLRLSLDAFHQKOTQPVID
YYAKKENLVNIPAEKPEEYTKVVKVST"
join(42526..42663,42914..42997,43149..43190,43341..43451,
43540..43731,43935..44032,44339..44516)
/note="gene_id:MX122.9
unknown protein"
/pseudo
/codon_start=1
/evidence=not experimental
/complement(46967..48880)
/note="gb|AAD25781.1
gene_id:MX122.10
strong similarity to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB09457.1"
/db_xref="GI:9758920"
/translation="MLLIFFSFDNRLSNIKNTAMTNOENQNDALVARLEAAESVIHRW
DGNDSRRHSSSSGNVRSSFSLSFDESPEATEFLSAVDEIIILLEDLSSENKPD
Query Match 3.9%; Score 171.2; DB 8; Length 80115;
Best Local Similarity 59.9%; Pred. No. 41e-24;
Matches 400; Conservative 0; Mismatches 148; Indels 120; Gaps 2;
QY 2827 TTAATAATTTAGTCCACCTGGCTCGGAAAGGGAACACAGTCACCGCTGATTAAGGA 2886
DB 40865 TTTCGAATTTGCAGGTCCACCTGGTTTCAGGAAAGGTACACAGTCTCCGGTCTATAAGGA 40806
QY 2887 TGAATTTTGTGCTGCCATTTAGCCACTGGTGATATGTTGAGGGCTGCAGTGGCTGCTAA 2946
DB 40805 TGAGTTTGTGTGTCATTTGCTACTGGTGACATGCTTAGAGTCTGCTGCTGCTAA 40746
QY 2947 AACTCCACTTGGGATTAAGGCTTAAGAAGCTATATGGAACAAGTAGTTTTTAAGAAACATAT 3006
DB 40745 GACTCCTCTGTGTGTAAGGCAAGGAAGCAATGGATA----- 40708
QY 3007 AGCAACAGAAATATATAACAGCAGGAGTGGTTCTTGTGATTCCTTTCTTCCTTAT 3066
DB 40707 ----- 40708
QY 3067 CTTCTAGGAGAGCTGTTTCGATGACTTGGTTGTTGGGATTTATGATGAAGCCATGAA 3126
DB 40707 -----AGGAGAGCTGTTTCTGATGACTTGGTTGTTGGTATCATGTGATGAAGCAATGAA 40653
QY 3127 GAAAACTTCATGCCAAGAAAGTTTATCCTTGTATGGTTTCCCTAGAACTGTTGTTCAAGC 3186
DB 40652 CAGACCCAAATGTGAGAAAGTTTCATTTCTGTATGGTTTCCCGGAGCTGTGACTCAAGC 40593
QY 3187 ACAGAAGGTGAGTCTCTTGGTCAATATGACCCGCTATATAAAGAGAGCTCTCTTTTGTAT 3246
DB 40592 AGAGAAGGTTTCTGTT-----CCTTTCTGTG 40566
```

```
QY 3247 TAGAGCTGCTATATAAATGCACAGTTTCTATCATTTGTATCACTTTTCTTACTAAAAAT 3306
DB 40565 GTTAGTTGTTGTGATCTCTCTGCTCTGCTTGTAAAGTTCTGACTTTTGTGTTATAAT 40506
QY 3307 GGTGAGCTTGATGAATGTTGGCCAAACAGGTACTAAGATTGACAAGTTCTTAAATTT 3366
DB 40505 CTTGTAGCTTGATGAGATGCTTAAATAGAGGGGAGCTCAGATAGATAAGGTGCTTAAATTT 40446
QY 3367 TGCATTTGATGATGCATATCTCGAAGAACGAATTACCGGCTCGTTGGATCCACCCATCAAG 3426
DB 40445 TGCATCGATGATTCGGTTCTCGAAGAAAGATTTCTGGAAGTGGATTCACCTTTCAAG 40386
QY 3427 TGGTAGATCTTATCATACAAAATTTGCTCTCTTAAGACTCTCGGACTTGGATGATGAAG 3486
DB 40385 TGGAAGAAGCTATCATACTAAATTCGCACCTCTCTAAAGTTCCAGGAGTCCGATGATGAAG 40326
QY 3487 TCATACCA 3494
DB 40325 AAAAACCA 40318
RESULT 6
RICKN57
LOCUS RICKN57 Rice mRNA for adenylate kinase (KN57 gene), linear PLN 02-FEB-1999
DEFINITION Rice mRNA for adenylate kinase (KN57 gene), partial sequence.
ACCESSION D10429
VERSION D10429.1 GI:218170
KEYWORDS adenylate kinase.
SOURCE Oryza sativa callus cDNA to mRNA.
ORGANISM Oryza sativa
REFERENCE 1 (sites)
AUTHORS Uchimiyama, H., Kidou, S., Shimazaki, T., Aotsuka, S., Takamatsu, S.,
Nishi, R., Hashimoto, H., Matsubayashi, Y., Kidou, N., Umeda, M. and
Kato, A.
Random sequencing of cDNA libraries reveals a variety of expressed
genes in cultured cells of rice (Oryza sativa L.)
Plant J. 2, 1005-1009 (1992)
2 (bases 1 to 387)
AUTHORS Kidou, S., Umeda, M., Kato, A. and Uchimiyama, H.
TITLE Plant cDNA homologue to rat insulinoma gene encoding ribosomal
protein S15
JOURNAL Nucleic Acids Res. 21 (8), 2013 (1993)
MEDLINE 93261842
PUBMED 8493114
REFERENCE 3 (bases 1 to 387)
AUTHORS Uchimiyama, H.
TITLE Direct Submission
JOURNAL Submitted (27-JAN-1992) Hirofumi Uchimiyama, Institute of Mol. &
Cell. Bioscience, The University of Tokyo, Department of Cellular
Function; 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113, Japan
Function; 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113, Japan
[ E-mail: huchimiyama@tansei.cc.u-tokyo.ac.jp ]
Tel: 03-3812-2111 (ex. 7844), Fax: 03-3812-2910
Submitted (27-Jan-1992) to DDBJ by:
Hirofumi Uchimiyama
Institute of Applied Microbiology
University of Tokyo
Bunkyo-ku, Tokyo 113
Japan
Phone: 03-3812-2111 x7844
Fax: 03-3812-2910.
Location/Qualifiers
FEATURES
source
1..387
/organism="Oryza sativa"
/db_xref="taxon:4530"
/tissue_type="callus"
BASE COUNT 108 a 71 c 104 g 104 t
ORIGIN
Query Match 3.8%; Score 164.4; DB 8; Length 387;
Best Local Similarity 74.0%; Pred. No. 6e-23;
```



FEATURES	source	Location/Qualifiers	
		1. .137348	
		/organism="Oryza sativa (japonica cultivar-group)"	
		/cultivar="Nipponbare"	
		/db_xref="taxon:39947"	
		/chromosome="1"	
		/clones="P0698A10"	
gene		/complement(join(202. .450,580. .654))	
		/gene="P0698A10.1"	
CDS		/complement(join(202. .450,580. .654))	
		/gene="P0698A10.1"	
		/note="hypothetical protein"	
		/codon_start=1	
		/protein_id="BAB92431.1"	
		/db_xref="GI:20804746"	
		/translation="MRYPRVQCAPVNSRSSMSDGRILHLPVGFAGFARSVVGWYGVRRP RGRPRRRGLNKGAAQAPRSYARDRVREISLCVTLRRRRRRPPYPAPTAATR RGRF"	
gene		/join(5591. .5603,5744. .5886)	
		/gene="P0698A10.2"	
CDS		/join(5591. .5603,5744. .5886)	
		/gene="P0698A10.2"	
		/note="hypothetical protein"	
		/codon_start=1	
		/protein_id="BAB92432.1"	
		/db_xref="GI:20804747"	
		/translation="MEKGEREGRGSGEGEGAHGALIGAVDSFRPVGEAGDSAVGR WRRWIGR"	
gene		/join(6553. .6815,6898. .7498)	
		/gene="P0698A10.3"	
CDS		/join(6553. .6815,6898. .7498)	
		/gene="P0698A10.3"	
		/note="contains ESTs C73695 (E20178), AU082323 (E20178)"	
		/codon_start=1	
		/product="putative myb-related protein"	
		/protein_id="BAB92433.1"	
		/db_xref="GI:20804748"	
		/translation="MGRSPCKEKAHTNKGAWTKEEDQRIIAYIKAHGSGCWRSIPKAA GLRCGSKLRWNYLRPDLCKGNFTDDDELIILKHALGNKWSLIAGLPGRTDN EKNYWNTHIKRLLRGIDPOTHRPVSAGSAAASGLTTTASTAAPFLAPAPPPQ QRLHNPVHAAPSNASFARSAASPESDGSHSSGSSDAPRCPLNDLDLDSMSL PSSPPTKPAASSTTASRHHHQQTICLYHLGVNRGSDVCSCKAAAPSPAGPAPRA FLRLESGOVI"	
gene		/complement(join(9381. .9448,10362. .10453,10723. .10803, 11573. .11701,12144. .12277,13341. .13677,16691. .16813, 16960. .17116,17804. .17874,17967. .18004))	
		/gene="P0698A10.4"	
CDS		/complement(join(9381. .9448,10362. .10453,10723. .10803, 11573. .11701,12144. .12277,13341. .13677,16691. .16813, 16960. .17116,17804. .17874,17967. .18004))	
		/gene="P0698A10.4"	
		/note="hypothetical protein"	
		/codon_start=1	
		/protein_id="BAB92434.1"	
		/db_xref="GI:20804749"	
		/translation="MRRERGLAEELDDAFTASTDIDPNTCEWPTSTSERPCVVQQA TYVYARQAEADPKPEEDAGREYACVLPVRIIOBYCNTNTSLOVSAYSVTRGHG LVQVKGVNATLWTITAPPFARNKNRVLQSGAFITGISTVCSSTCHGTQLMSLSQV CCRCNGLHLPVWKODGQGFHALFPVASTRSVAHSYTVYVILLQOVKHSRISLS KWTVANDSOVGSNTLKNVRYQLOLAKAKRDPFACRGGGRFNACRERQQLAI SMDDPPDCPVSIIRHKNMPKVQTIPLNLSRSDGRRFRFQGLRIPDGGERLLIQ INMEKGNKKNKGVGRVPPFTLIQPRHEDHTDRSPVLATKRNQOQOQKAQD MGFGKTWAKPEKNSER"	
gene		/join(19424. .19431,19465. .19629,20505. .20545,20966. .21079, 21836. .21852)	
		/gene="P0698A10.5"	
CDS		/join(19424. .19431,19465. .19629,20505. .20545,20966. .21079, 21836. .21852)	
		/gene="P0698A10.5"	
		/note="hypothetical protein"	
		/codon_start=1	
		/protein_id="BAB92435.1"	
		/db_xref="GI:20804750"	
		/translation="MAVKKINFDKLCIGRTLIVINGSTLHRSPPRPIKSDDDTATRLVLP RLARLYNICDRITCCFWYDIVVAHAENDHVELREIRIILQRLDQQLCIGFSWRRD TFVSLPALEQR"	
gene		/join(23333. .24646,25086. .25210,25302. .25443,25546. .25806)	
		/gene="P0698A10.6"	
CDS		/join(23333. .24646,25086. .25210,25302. .25443,25546. .25806)	
		/gene="P0698A10.6"	
		/note="contains ESTs AU093915 (E1276), AU162319 (E50301)"	
		/unknown protein"	
		/codon_start=1	
		/protein_id="BAB92436.1"	
		/db_xref="GI:20804751"	
		/translation="MAEEFQPEAAPAAVAATTEVAVAEKAPVEAEKEKKVEETPAVE AEAKEEKDEAAAAAGQDEAGAEIGTGSFKEESNLVADLPDEKKALDFKQLIAA ALACENFLUPPPPPPKAKVEAAVEETKABETKABEEKPAEPAKEEPAEVAANAA APACETEPEKPAEASSEAEKPEKAAEADPEKAEKSAEAPAEAEKPAEPEPEEK TVVTEEAATKTVEAIEETVVVAAAAAATAEAAAAPEVEQAAAAPEVLTWGV LVGDDERTVTLLKFLRAREFKVKAMALRSVLRKRFGIESLDDLALPELDSV VFYRGADRGHPVCNVYGEFODKLYEKAFGDEEKRERFLKWRIQLLERGILSQLD SPSGICSMYQVTDLKNSPMLGKHRAVTRQAVALLQDNYPEFIKKFINVPMWYLA NKMSPFLQRTDKSKFIFASPAKSAETLFRITAEQVPVQFGLPKEDDPETTSDAV TELTIKPSKETVEIPVTENSTIGMELRVLGEVSYGAETPDABEGVTVLVQKTRKV PANEEPIMKSGFKVGPCKIVLTINNPAKSKKKLLYRSKVKSTSES"	
gene		/join(27824. .27925,28000. .28127,28637. .28780,28932. .29064)	
		/gene="P0698A10.7"	
CDS		/join(27824. .27925,28000. .28127,28637. .28780,28932. .29064)	
		/gene="P0698A10.7"	
		/note="hypothetical protein"	
		/codon_start=1	
		/protein_id="BAB92437.1"	
		/db_xref="GI:20804752"	
		/translation="MSGSPFAAAAAAIALASWANKLERDRVATRRWEGPAIAKPATI WGPPIRRRTATAVQTAVRTPDSPRRRHGRHSATRRMRLHSLVLPANACRIAA VNHARRIGNOKTRTPTLPRCPHAIIRRVWFMLRAAAAAASVLRQFNTMPDLNCQRAY EQIQCVV"	
gene		/complement(join(31579. .31693,31909. .32131,32235. .32469, 32585. .32716,32809. .33174))	
		/gene="P0698A10.8"	
CDS		/complement(join(31579. .31693,31909. .32131,32235. .32469, 32585. .32716,32809. .33174))	
		/gene="P0698A10.8"	
		/codon_start=1	
		/product="putative DNA polymerase I"	
		/protein_id="BAB92438.1"	
		/db_xref="GI:20804753"	
		/translation="MPPFAVSTAPVAVPVVASSSSAAAAAPRRSLRLURVTYASVSS SSPSTFAVISIPPPPARSHRQLVGRDGAAPAKPRVFFLDVNPCLFCFGSQRSLGA FARWALFFAHVSLRDPVAVLDGEGNEYRRLLPSYKAHRPGPGTGADSRVIDVL RCGNPVVRDGYEADDVATLVEQVLQKGFVWIGSPDKQFKQLISDVQVMPIDE IGRWSFYTLRHVVAQYKCDPTADLSLRCFIGEADGIPGIOHLVFGFRKTAVKLLKK HGSLENLNTAAVTVGKDYAQDALVKHADVLKKNYEVLSLKRDVKVQLDDRLWLSRD SNDSSVLSDFLKFNEGQNR"	
gene		/complement(34177. .34410)	
		/gene="P0698A10.9"	
CDS		/complement(34177. .34410)	
		/gene="P0698A10.9"	
		/note="hypothetical protein"	
		/codon_start=1	
		/protein_id="BAB92439.1"	
		/db_xref="GI:20804754"	
		/translation="MTSUTGRAFRDDNGFFIFATELARSMA DTNTDLHVTSPTPAARR SWHPFSLLLTASRRMPSGKKQKRRKLEI"	
gene		/complement(join(35312. .35473,36485. .36646))	
		/gene="P0698A10.10"	
CDS		/complement(join(35312. .35473,36485. .36646))	
		/gene="P0698A10.10"	
		/note="hypothetical protein"	
		/codon_start=1	
		/protein_id="BAB92440.1"	
		/db_xref="GI:20804755"	
		/translation="NEHGRGEEFGRETGWEDVAPADTCRSGYPVWAGFFFWAVST NRRLTSDILOESGIPVPRIQFLCTOOLASKRNLFSENRSTRDPFKNTPSDSKSYND EYRL"	

```
gene      join(37075..38148,39504..39906,40314..40415,40513..40736)
          /gene="P0698A10.11"
CDS       join(37075..38148,39504..39906,40314..40415,40513..40736)
          /gene="P0698A10.11"
          /note="contains ESTs AU101336(E11055),AU101337(E11055)"
          /codon_start=1

Query Match      2.7%; Score 115.6; DB 8; Length 137348;
Best Local Similarity 76.3%; Pred. No. 8.3e-13;
Matches 142; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 893 TTTCAACGATCATATAACGAGAAAGCTCATTAGACATTTATTCTAGATATTTATAAT 952
      |||
Db 136817 TTTTAACTCACCACCAACGAGAAACTTTATTAGCACAATAATTAAGTATTTATAAT 136758

QY 953 TATAAATTTGAAAAAATATTTTATTGAAATTTTAAACAAATGTATGCATAAATTTT 1012
      |||
Db 136757 TATAAATTTGAAAAAATATTTTATTGATTTTAAATAAATTTTATATAGATTTT 136698

QY 1013 TTAACACACACCAATTAACCTTTAAAGCATCTCTAATAGGAACGAGAACTTAA 1072
      |||
Db 136697 TTAATAACATATCATTTAACTTTGAAAGCATCTAATACGGAACGATAAAATTTAA 136638

QY 1073 GATTCA 1078
      |||
Db 136637 GTTTGA 136632

RESULT 9
AY079340 AV079340 772 bp mRNA linear PLN 18-APR-2002
LOCUS Arabidopsis thaliana putative adenylylate kinase (At5g63400) mRNA,
complete cds.
ACCESSION AY079340.1 GI:19310680
VERSION FLI CDNA.
KEYWORDS Arabidopsis thaliana.
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 772)
Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Egtu,P.,
Lee,J.M., Toriumi,M., Yu,G., Brooks,S., Chao,Q., Chen,H.,
Karlin-Neumann,G., Kim,C., Lam,B., Miranda,M., Nguyen,M.,
Palm,C.J., Shinn,P., Southwick.A., Davis,R.W., Ecker,J.R. and
Theologis,A.
Aradidopsis Open Reading Frame (ORF) Clones
Unpublished
2 (bases 1 to 772)
Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
Direct Submission
Submitted (19-FEB-2002) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
COMMENT The RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN
Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai, J.,
Hayashizaki,Y. and Shinozaki,K.

TITLE The Salk, Stanford, PGEC (SSP) Consortium members constructed and
JOURNAL sequenced the PUNI (ORF) clones using the RAFL cDNAs: Yamada,K.,
REFERENCE Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M.,
AUTHORS Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C.,
Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H.,
Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J.,
Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
Direct Submission
Submitted (19-FEB-2002) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
COMMENT The RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN
Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai, J.,
Hayashizaki,Y. and Shinozaki,K.
```

```
MEYERS,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P.,
Southwick.A., Davis,R.W., Ecker,J.R. and Theologis,A.
Yamada,K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally
to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP
/PGEC) contributed equally to this work as PIs.

FEATURES
Location/Qualifiers
1..772
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="5"
/clone="U09821"
/note="This clone is in PUNI 51."
ecotype: Columbia
1..772
/gene="At5g63400"
1..741
/gene="At5g63400"
/codon_start=1
/evidence=experimental
/product="putative adenylylate kinase"
/protein_id="AA185071.1"
/db_xref="GI:19310681"
/translation="MATGGAADLEVDVQVLMSELLRLKCSQPKRLIFIGPPGS
GKGTSVPVKQDEYCLHSTGDMRAAVASTPLGVKAKAMEKGLVDDLVVGIID
EAMNPKCKQKGFILDGFPRTVTOAEKIDMLKRRGTEIDKVLNFAIDAIEERITGR
WIHPSSGSYHTKFPAPKTPGVDDITGEPLIQKDDNADVLKSLAAPHQSOTQPVIDY
YAKKAVLTNIOAEKAPQEVTSVKALS"
742..772
/gene="At5g63400"
3'UTR
BASE COUNT 221 a 175 c 189 g 187 t
ORIGIN
Query Match 2.6%; Score 112.8; DB 8; Length 772;
Best Local Similarity 76.7%; Pred. No. 1.9e-12;
Matches 138; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 3308 GTGCAGCTTCATGAATTTGGCCAAACAAAGGTACTTAAGATTGACAAAGTTCTTAAATTT 3367
      |||
Db 379 GAGAAGCTCGATGAGATGCTTAAGAGCGGAGCACTGAAATTTGACAAAGTTCTCAACTTT 438

QY 3368 GCAATTGATGATGCATATCTGGAAGAACCAANTACCGTTCGTTGGATCCACCATCAAGT 3427
      |||
Db 439 GCTATTGATGACCAATCTTTGGAGAAAGAATAACCGGCGATGGATCCACCATCGAGT 498

QY 3428 GGTAGATCTTTATCATACAAAATTTGCTCTCTCTTAAGACTCTCGGACTTGATGATGTAAGT 3487
      |||
Db 499 GGCAGGAGTTACCACACCAAAATTTGCTCTCTCCAAACCCCTGGAGTTGATGATATTACT 558

RESULT 10
AY085188 Arabidopsis thaliana clone 13667 mRNA linear PLN 21-JUN-2002
LOCUS Arabidopsis thaliana clone 13667 mRNA, complete sequence.
DEFINITION Arabidopsis thaliana clone 13667 mRNA, complete sequence.
ACCESSION AY085188.1 GI:21403898
VERSION FLI CDNA.
KEYWORDS Arabidopsis thaliana
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 904)
Haas,B.J., Volfovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,
Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
Full-length messenger RNA sequences greatly improve genome
annotation
Genome Biol. (2002) In press
2 (bases 1 to 904)
Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
Full-length cDNA from Arabidopsis thaliana
Unpublished
```



Db 393 GAGAACTCGATGAGATGCTTAAGAGCGGAGGAACGTGAATTCACAAAGTTCTCAACTTT 452  
 QY 3368 GCAATTGATGATCAATCTACTGGGAAGACGAATTTACGGTCTGGTGGATCCACCCATCAAGT 3427  
 Db 453 GCTATTGATGACGCAATCTTTGGAGGAAGAATAACCGGGCGATGGATCCACCCATCGAGT 512  
 QY 3428 GGTAGATCTTATCATCAAAATTTGCTCTCTCTTAAGACTCCTGGACTTGGATGATCAAGT 3487  
 Db 513 GCGAGGAGTTACCACACCAAAATTTGCTCTCTCCCAAAACCCCTGGAGTTGATGATATTACT 572

RESULT 12  
 LOCUS AP005594 124815 bp DNA linear HTG 24-JUL-2002  
 DEFINITION Oryza sativa (japonica cultivar-group) chromosome 9 clone P0701E06,  
 \*\*\* SEQUENCING IN PROGRESS \*\*\*, in ordered pieces.

ACCESSION AP005594  
 VERSION AP005594.1 GI:21952963  
 KEYWORDS HTG: HTGS PHASE2.  
 SOURCE Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,  
 clone:P0701E06.

ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1

AUTHORS Sasaki,T., Matsumoto,T., Hattori,M., Sasaki,Y. and Katayose,Y.  
 TITLE Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, PAC  
 clone:P0701E06

JOURNAL Published Only in Database (2002)

REFERENCE 2 (bases 1 to 124815)

AUTHORS Sasaki,T., Matsumoto,T., Hattori,M., Sasaki,Y. and Katayose,Y.

TITLE Direct Submission

JOURNAL Submitted (23-JUL-2002) Takuji Sasaki, National Institute of

Agrobiological Sciences, Rice Genome Research Program; Kannondai

2-1-2, Tsukuba, Ibaraki 305-8602, Japan

(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/.

TEL:81-298-38-7441, Fax:81-298-38-7468)

NOTE: It currently consists of 1 contigs. Gaps between the contigs

are represented as runs of N. The order of the pieces is believed

to be correct as given, however the sizes of the gaps between them

are based on estimates that have provided by the submitter. This

sequence will be replaced by the finished sequence as soon as it is

available and the accession number will be preserved.

\* NOTE: This is a 'working draft' sequence.

\* This sequence will be replaced

\* by the finished sequence as soon as it is available and

\* the accession number will be preserved.

FEATURES

source

1. .124815

Location/Qualifiers

/organism="Oryza sativa (japonica cultivar-group)"

/cultivar="Nipponbare"

/db xref="taxon:39947"

/chromosomes="9"

/clone="P0701E06"

BASE COUNT 37055 a 25766 c 25390 g 36503 t 101 others

ORIGIN

Query Match 2.5%; Score 109.2; DB 2; Length 124815;

Best Local Similarity 63.2%; Pred. No. 1.6e-11;

Matches 168; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 820 TAAGAACGATGAGTTGGGAACACAGAGAAACACAGCGCTTAAGGCTTTCTTGATCTCTTA 879

Db 56107 TTATAAGTATGTTGTAGGATTAATCAATTAAGATGACTAATGCTCGGTTCAATCACCTC 56166

QY 880 GTTCGAGTTGATTTTCACGCGATGATAACGAGAAAGCTCATTAGCACATTATTACTT 939

Db 56167 ATTTGGAGATAGAAATTTCAACGCGCACAAAACGAGAAATCTCATTAGCACATGATCAAT 56226

QY 940 AGATATTATTAATTAATAAATTTGAAAAAATATTTATTTGAAATTTTAAACATGATG 999

Db 56227 AAGTGTAAATTAATTAATAAATTTGAAAAATGGAATTTATTTGATATTTTAACTACTTCTA 56286

QY 1000 CATAAATTATTTTAAAAACACACCAATTTAAACCTTTAAAAAGCATCTTAATAGAAA 1059  
 Db 56287 TATAGATTTTAAAAAGAACACATCATTTAACTGTGTGAAAAAGTATGCTAATGAAAA 56346  
 QY 1060 CGAGGAAGTTAAAGATTACCCGAAAT 1085  
 Db 56347 TGAGAGAGTTGAAGTTTACTAGAGT 56372

RESULT 13

LOCUS AP003261/c

DEFINITION

Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,

PAC clone:P0471B04.

ACCESSION AP003261

VERSION AP003261.3 GI:2202658

KEYWORDS

SOURCE

ORGANISM

Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,  
 clone:P0471B04.

Eukaryota; Viridiplantae; Streptophyta;  
 Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales;  
 Poaceae;  
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1

AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.

TITLE Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC  
 clone:P0471B04

JOURNAL Published Only in Database (2001)

REFERENCE 2 (bases 1 to 138025)

AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.

TITLE Direct Submission

JOURNAL Submitted (19-FEB-2001) Takuji Sasaki, National Institute of

Agrobiological Sciences, Rice Genome Research Program; Kannondai

2-1-2, Tsukuba, Ibaraki 305-8602, Japan

(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/.

TEL:81-298-38-7441, Fax:81-298-38-7468)

On Aug 9, 2002 this sequence version replaced gi:21202837.

Genes were predicted from the integrated results of the following:

GENSCAN1.0, BLASTX2.0, BLASTY2.0 as well as SplicePredictor

(October 1998 version). The genomic sequence was searched against

NCBI NonRedundant Protein database, nr

(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at

RGF. Protein homologies of the coding regions were searched against

NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent

the identified cDNA sequences using BLASTN 2.0 with the

corresponding DBJ accession no. and RGP clone ID.

A gene with identity or significant homology to a protein is

classified based on the protein name to indicate the homology level

such as same name, 'putative-' and '-like protein'. A gene without

significant homology to any protein but with EST homology (covering

almost the entire length of partial sequence) is classified as an

'unknown' protein. A gene predicted with a gene prediction program

is classified as a 'hypothetical' protein.

The orientation of the sequence is from SP6 to T7 of the PAC clone.

This sequence of P0471B04 clone has an overlap with P0698A10 clone

(DBJ : AP003297) at the position 1 to 24,438 of 5' end and an

overlap with P0018C10 clone (DBJ: AP003227) at the position 48,801

to 138,025 of 3' end. Detailed information on overlap and assembly

quality together with annotation of this entry is available at

http://rgp.dna.affrc.go.jp/genomeSeq.html.

FEATURES

source

1. .138025

Location/Qualifiers

/organism="Oryza sativa (japonica cultivar-group)"

/cultivar="Nipponbare"

/db xref="taxon:39947"

/chromosomes="1"

/clone="P0471B04"

join(9533..9715,9887..10019,10850..10988,11365..11401,

13640..13931,14823..14955,15564..15576)

/gene="P0471B04.1"

join(9533..9715,9887..10019,10850..10988,11365..11401,

13640..13931,14823..14955,15564..15576)

gene

CDS







```

Db 23907 TTTTAACTCAGCAAAACAGAGAAACTTATTAGCACATAATTAATTAAGTATTAAAT 23848
Qy 953 TATAAAGCTGAAAAAATAATTTATTTGAATTTTTTAAACAATGTATGCAATAATTTATTT 1012
Db 23847 TATAAAGCTGAAAAAATAATTTATTTGAATTTTTTAAACAATGTATTTATATAGATTTTTT 23788
Qy 1013 TTAATAACACACCAATTTTACCTTTTAAAGAGCATCTTAATAGGAACGAGGAAGTTAAA 1072
Db 23787 TTAATAACATATCATTTTACAGTTTGAAGAGCATCTAATACGGAACGATAAAATTTAAA 23728
Qy 1073 GATTCA 1078
Db 23727 GTTTGA 23722

RESULT 14
CNS07YYPX 169663 bp DNA linear HTG 27-JUL-2002
LOCUS Oryza sativa chromosome 12 clone OJ1618_C05, *** SEQUENCING IN
DEFINITION PROGRESS ***, in ordered pieces.
ACCESSION AL713952
VERSION AL713952.2 GI:22003292
KEYWORDS HTG; HTGS_PHASE2; HTGS_ACTIVEFIN.
SOURCE Oryza sativa
ORGANISM Oryza sativa
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 169663)
Choise.N., Orjeda,G., Cattolico,L., Demange,N., Wincker,P.,
Segurens,B., Pellatier,E., Scarpelli,C., Salanoubat,M.,
Weissenbach,J. and Quetier,F.
Oryza sativa chromosome 12 sequencing
Unpublished
2 (bases 1 to 169663)
Genoscope.
Direct Submission
Submitted (26-JUL-2002) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
On Jul 29, 2002 this sequence version replaced gi:19715885.
IMPORTANT: This sequence is unfinished and does not necessarily
represent the correct sequence.
Work on the sequence is in progress and the release of this data is
based on the understanding that the sequence may change as work
continue. The sequencing may be contaminated with foreign sequence
from E.coli, yeast, vector, phage, etc.
The nucleotide sequence of this BAC clone was generated by
combining Monsanto and Genoscope sequencing data.
* NOTE: This is a working draft sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
Location/Qualifiers
source
1. 169663
/organism="Oryza sativa"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:4530"
/chromosomes="12"
/clone="OJ1618_C05"
/clone_lib="Monsanto"
BASE COUNT 46256 a 37378 c 36384 g 49645 t
ORIGIN

Query Match 2.5%; Score 109.2; DB 2; Length 169663;
Best Local Similarity 68.8%; Pred. No. 1.7e-11;
Matches 150; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

Qy 861 AAGGCTTCTGATCCTCTAGTCGAGGTGATTTTCAACGCATGATAACGAGAAAGCT 920
Db 103185 AAGCACTAGGCTTCGGCTTTGATAGAAAAATTTTAAACGACGCAAAACAGAAAGCT 103244

```

```

Qy 921 CATTAGCACATTTACTTAGATATTTATTAATAATAAACTTCAAAAAATATTTATTTGA 980
Db 103245 CATTAGCACATAATGATTAAAGTATTAAATAATAATAAAATGGAATTTATTTGA 103304
Qy 981 ATTTTTTAAACAATGATGCATAAAATTTATTTTAAACAACACCAAAATTTAAACCCCTTTAA 1040
Db 103305 TTTTTTAAACAACCTTCTATATAGAAATTTTCTTAAACAACACATCATTTAAACAGTTTAA 103364
Qy 1041 AAAGCATCTCTAATAGAAACGAGGAAGTTAAAGATTCA 1078
Db 103365 AAAGCGTGCTTAACAGAAAACGAGGAGTTTGAAGTTTAA 103402

RESULT 15
OSJN00015 126659 bp DNA linear PLN 17-JUL-2002
LOCUS Oryza sativa genomic DNA, chromosome 4, BAC clone: OSJNB0079B02,
DEFINITION complete sequence.
ACCESSION AL606455
VERSION AL606455.2 GI:21740404
KEYWORDS HTG
SOURCE Oryza sativa.
ORGANISM Oryza sativa
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1
Han,B., Feng,Q., Huang,Y.C., Li,Y., Zhu,J.J., Zhao,Q., Hu,X.,
Liu,Y.L., Mu,J., Yu,Z., Chen,L., Fan,D.L., Weng,Q.J., Zhang,L.,
Lu,Y.Q., Yu,S.L., Liu,X.H., Lu,T.T., Zhang,Y.J., Lu,Y., Li,C.,
Li,T., Zhang,Y., Hu,H., Jia,P.X., Qian,Y.M., Ying,K., Zhou,B.,
Chen,Z.H., Hao,P., Zhang,L., Wu,M., Zhang,R.Q., Guan,J.P., Fu,G.,
Wang,S.Y., Ren,S.X., Lv,G., Lin,W., Gu,W.Q., Zhu,G.F., Tu,Y.F.,
Jia,J., Yin,H.F., Zhang,Y., Cai,Z., Chen,J., Kang,H., Chen,X.Y.,
Shao,C.Y., Sun,Y., Hu,Q.P., Zhang,X.L., Zhang,W., Wang,L.J.,
Ding,C.W., Sheng,H.H., Gu,Q.L., Chen,S.T., Ni,L., Zhu,F.H. and
Hong,G.F.
Direct Submission
Submitted (27-JUN-2002) Han Bin, National Center for Gene Research,
Chinese Academy of sciences, 500# Cao Bao Road, Shanghai 200233,
CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests:
bhan@ncgr.ac.cn
Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC
clone: OSJNB0079B02.
On Jul 12, 2002 this sequence version replaced gi:15552693.
Web site: http://www.ncgr.ac.cn
----- Summary Statistics
Assembly program: phrap
-----
NOTE: This is a complete sequence.
Genes were identified by a combination of several methods: Gene
prediction programs including Fgenesh (http://www.softberry.com/),
genscan (http://cgr-081.mit.edu/GENSCAN.html), GeneMarkHM
(http://genemark.biology.gatech.edu/Genemark/), tRNAscan-SE (Sean
Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/), searches of the
complete sequence against NCBI none redundant protein database (nr)
(http://ncbi.nlm.nih.gov/blast/db) and the EST database at NCGR.
Location/Qualifiers
source
1. 126659
/organism="Oryza sativa"
/variety="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:4530"
/chromosomes="4"
/clone="OSJNB0079B02"
/clone_lib="CUGI-OSJNBb"
BASE COUNT 35469 a 27283 c 28042 g 35865 t
ORIGIN

Query Match 2.5%; Score 108.8; DB 8; Length 126659;
Best Local Similarity 55.9%; Pred. No. 2e-11;
Matches 227; Conservative 0; Mismatches 177; Indels 2; Gaps 1;

```



GenCore version 5.1.1.4 p5 4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2003, 09:09:00 ; Search time 809.479 Seconds  
(without alignments)  
12076.811 Million cell updates/sec

Title: US-09-802-937-1

Perfect score: 4341

Sequence: 1 cgcgaggaagtaattagg.....gtatatcttcattgtttt 4341

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
3	1440	33.2	1440	21	AAA07482
4	1440	33.2	1440	21	AAA10961
5	115.8	20.7	809	21	AAC48619
6	115.2	2.7	482	21	AAC48110
7	113.2	2.6	489	21	AAC38309
8	112.8	2.6	904	21	AAC47002
9	112.8	2.6	981	24	ABN98503

10	102.6	2.4	608	21	AAC49033
11	102.6	2.4	611	21	AAC35687
12	102.6	2.4	1017	21	AAC49022
13	102.6	2.4	1023	21	AAC37085
14	102.2	2.4	295	24	ABL72417
15	101.8	2.3	894	24	ABN98716
16	100	2.3	422	24	ABL49953
17	97.4	2.2	423	24	ABL49955
18	97.4	2.2	424	24	ABL49954
19	66.6	1.5	5875	24	ABL32289
20	64.8	1.5	6048	24	ABO67002
21	64	1.5	9542	24	ABL32357
22	63	1.5	6095	22	RA546310
23	63	1.5	6095	24	ABL70150
24	63	1.5	6095	24	ABL32361
25	63	1.5	6095	24	ABL34475
26	62.6	1.4	8237	22	RA546802
27	62	1.4	721	21	AAF14639
28	61.6	1.4	645	23	RA553274
29	61.4	1.4	682	21	AAC44208
30	61.4	1.4	883	22	RA115210
31	61.4	1.4	15548	24	ABL34155
32	61.2	1.4	7165	24	ABL70222
33	61.2	1.4	7165	24	ABL32751
34	61.2	1.4	7165	24	ABK31267
35	61	1.4	736	21	AAF07686
36	60.8	1.4	3587	23	ABL05676
37	60.6	1.4	38342	22	RA546746
38	60.6	1.4	38342	24	ABK31507
39	60.4	1.4	6079	24	ABL32421
40	60.4	1.4	7442	22	RA546686
41	59	1.4	38342	22	RA546746
42	59	1.4	38342	24	ABK31507
43	58.8	1.4	13202	24	ABL33485
44	58.6	1.3	11836	22	RA545395
45	58.6	1.3	11836	24	ABK28240

#### ALIGNMENTS

#### RESULT 1

AAA07481  
ID AAA07481 standard; DNA; 4341 BP.

XX AAA07481;

DT 17-JUL-2000 (first entry)

XX Rice adenylate kinase DNA sequence fragment #1.

DE Rice; adenylate kinase; promoter; structural gene expression;

XX plant gene expression; herbicide resistance; stress resistance;

KW light resistance; antibacterial production; flowering induction;

KW insecticidal substance production; ss.

XX Oryza sativa.

OS Oryza sativa.

XX WO200015612-A1.

PN 23-MAR-2000.

XX 08-SEP-1999; 99WO-JP04847.

XX 10-SEP-1998; 98WO-JP04088.

PR (NISC) NISSAN CHEM IND LTD.

XX Uchimiya H, Azai S, Fushimi T, Tagawa M, Fukuzawa H;

XX WPI; 2000-271448/23.

XX Rice adenylate kinase gene-originated DNA fragment with promoter

PT

PT function in plants for controlling expression of desired foreign  
 PT structural genes, useful for herbicide resistance and in sucrose  
 XX induction in recombinant plants -

Claim 1; Page 25-27; 33pp; Japanese.

XX This sequence represents a fragment of the rice adenylate kinase gene.  
 CC The invention relates to a DNA fragment (such as this sequence), that has  
 CC a regulatory function (preferably acting as a promoter), that has  
 CC desired foreign structural genes in a plant when introduced into a  
 CC plant body or plant cell. The DNA fragment can be inserted into a plant  
 CC to control the expression of desired foreign structural genes, useful for  
 CC inducing in the plant e.g. herbicide, stress and light resistance,  
 CC production of antibacterial and insecticidal substances, sucrose  
 CC induction in recombinant plants, and in promoting flowering and crop  
 CC yield. The DNA fragment has a regulatory function on the expression of  
 CC desired structural genes in a plant, particularly with control of the  
 CC expression in specific tissues and locations to provide safe chemicals.

XX Sequence 4341 BP; 1284 A; 758 C; 880 G; 1419 T; 0 other;

Query Match 100.0%; Score 4341; DB 21; Length 4341;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 4341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCAGGAAGATTAAATAGTGGACACACCAACCCCTGTGGTGGTGACGCCCTGTGTT 60

Db 1 CTGCAGGAAGATTAAATAGTGGACACACCAACCCCTGTGGTGGTGACGCCCTGTGTT 60

Qy 61 AATCAACTGGGGTTCGTTGGACATGTTTTTTCAGGAAATTAAGCAAGAAATTAAG 120

Db 61 AATCAACTGGGGTTCGTTGGACATGTTTTTTCAGGAAATTAAGCAAGAAATTAAG 120

Qy 121 AAGAACTCTCAAGCTGACATGAGAAACGTAATCCATGAGCGGAATTCAGTCGTC 180

Db 121 AAGAACTCTCAAGCTGACATGAGAAACGTAATCCATGAGCGGAATTCAGTCGTC 180

Qy 181 TCTTGTAACCTACCTGTTAGATACATAGACAGTCCCAACGTTTGATGGCTCCTATTGG 240

Db 181 TCTTGTAACCTACCTGTTAGATACATAGACAGTCCCAACGTTTGATGGCTCCTATTGG 240

Qy 241 CTGCTGTGATACCTGACCTGTCACAAAGCATCAATTCGTTCTGGAGTATCTTAATA 300

Db 241 CTGCTGTGATACCTGACCTGTCACAAAGCATCAATTCGTTCTGGAGTATCTTAATA 300

Qy 301 CCGAAAAACCCCAAGATTAATCTTCCACTCAGGGTAATGTGCTGAACATATGCAATG 360

Db 301 CCGAAAAACCCCAAGATTAATCTTCCACTCAGGGTAATGTGCTGAACATATGCAATG 360

Qy 361 AATACAAATTCGCAAAATATCATGTTATCTATCTTGCTCAAAATTTGAGTCCAA 420

Db 361 AATACAAATTCGCAAAATATCATGTTATCTATCTTGCTCAAAATTTGAGTCCAA 420

Qy 421 CTGAGACTGCAATACGATTTCTTTTCAAAAGAAATTAATTAATTTTTCATGAAA 480

Db 421 CTGAGACTGCAATACGATTTCTTTTCAAAAGAAATTAATTAATTTTTCATGAAA 480

Qy 481 CGCAATTCACCGTTCAGAAATATGCTGTCAATAAATAGTAGTCTAGTGCAGAAACAAA 540

Db 481 CGCAATTCACCGTTCAGAAATATGCTGTCAATAAATAGTAGTCTAGTGCAGAAACAAA 540

Qy 541 ATTAATATACATAAAAGAGGTTGTTAATTAACAAACCATGTTTCGTACTACAACTCT 600

Db 541 ATTAATATACATAAAAGAGGTTGTTAATTAACAAACCATGTTTCGTACTACAACTCT 600

Qy 601 AATTTGTAAATCTTATTTTCAGTCACAAAATTCCAATTTCCAATTAAGAAAAATAAAGCT 660

Db 601 AATTTGTAAATCTTATTTTCAGTCACAAAATTCCAATTTCCAATTAAGAAAAATAAAGCT 660

Qy 661 AGACGGCTAAGCCGCCCTCTAAGCTAAGTTCGAGAGGTGAAGTACGCAGAAAAATA 720

Db 661 AGACGGCTAAGCCGCCCTCTAAGCTAAGTTCGAGAGGTGAAGTACGCAGAAAAATA 720

Qy 721 TGATGGTTTATTAATATGATTTTTTTTAAATAAATCTTTACATAAATTTCTTTAGGAAACA 780

Db 721 TGATGGTTTATTAATATGATTTTTTTTAAATAAATCTTTACATAAATTTCTTTAGGAAACA 780

Qy 781 TATCATTTAATGTTTGAANAACGTGCACATAGAAAACCTTAAGAACGATGAGTTGGGAAA 840

Db 781 TATCATTTAATGTTTGAANAACGTGCACATAGAAAACCTTAAGAACGATGAGTTGGGAAA 840

Qy 841 CAAGAGAAAAACACAGCTTTAAGGCTTTCTTGATCCTCTAGTTGGAGGTTGATTTTCAAAC 900

Db 841 CAAGAGAAAAACACAGCTTTAAGGCTTTCTTGATCCTCTAGTTGGAGGTTGATTTTCAAAC 900

Qy 901 GCATGATAAAGGAGAAAGCTCATAGACATATTACTTTAGATATTATTAATTTATAACT 960

Db 901 GCATGATAAAGGAGAAAGCTCATAGACATATTACTTTAGATATTATTAATTTATAACT 960

Qy 961 TGAATAAATAATTTATTTGAATTTTTTAAACAATGTATGCATAAATTTATTTTAAAAAC 1020

Db 961 TGAATAAATAATTTATTTGAATTTTTTAAACAATGTATGCATAAATTTATTTTAAAAAC 1020

Qy 1021 ACACCAATTTAAACCTTTAAAAAGCATCTTAATAGAAAACGAGGAAAGTTAAAGATTCAAC 1080

Db 1021 ACACCAATTTAAACCTTTAAAAAGCATCTTAATAGAAAACGAGGAAAGTTAAAGATTCAAC 1080

Qy 1081 GAAGTGTTCGATAATGAAAAATGSGGTGGGATTAGCAATTTGGTAAATGAATCAGGGTTAG 1140

Db 1081 GAAGTGTTCGATAATGAAAAATGSGGTGGGATTAGCAATTTGGTAAATGAATCAGGGTTAG 1140

Qy 1141 GATTAAATATTAAATGAAAGAGGGAGAAATGAATGTTTAGAGTTTAAATGTGTCTTTTGG 1200

Db 1141 GATTAAATATTAAATGAAAGAGGGAGAAATGAATGTTTAGAGTTTAAATGTGTCTTTTGG 1200

Qy 1201 GTGGGTAGAAAAATTTATTTGCCATACACTCCCGAGAGCGGTGGTCTTCCGTGGGCGAGA 1260

Db 1201 GTGGGTAGAAAAATTTATTTGCCATACACTCCCGAGAGCGGTGGTCTTCCGTGGGCGAGA 1260

Qy 1261 AGCGTCTTTTTTCGTTGGAAAAAAAATCTGCTTTAAAAAGAAAAACAGAAAGAGCCAGCTTT 1320

Db 1261 AGCGTCTTTTTTCGTTGGAAAAAAAATCTGCTTTAAAAAGAAAAACAGAAAGAGCCAGCTTT 1320

Qy 1321 GTTGTACACGCTCTCACAGAGAAAGAAAGCCCAACCCACCTTAAACCTCTCGATCC 1380

Db 1321 GTTGTACACGCTCTCACAGAGAAAGAAAGCCCAACCCACCTTAAACCTCTCGATCC 1380

Qy 1381 GACCGAGACTCTCTCCATTTTCCAGCGGCGACCGGGAGAGACACGCGAGTTCGTTCGAC 1440

Db 1381 GACCGAGACTCTCTCCATTTTCCAGCGGCGACCGGGAGAGACACGCGAGTTCGTTCGAC 1440

Qy 1441 GAACAGGCTTAGTGAGTAGTTGTTGCTCGGAGATGGCGGGGAACTGGAGAGAGTGCC 1500

Db 1441 GAACAGGCTTAGTGAGTAGTTGTTGCTCGGAGATGGCGGGGAACTGGAGAGAGTGCC 1500

Qy 1501 GTCGATGGAGCTCATACAGAGCTGCTCCGCGCATGAAGTCAGCTCCAAGCCGACAA 1560

Db 1501 GTCGATGGAGCTCATACAGAGCTGCTCCGCGCATGAAGTCAGCTCCAAGCCGACAA 1560

Qy 1561 GCSCGTCATCTCGTCGTAACGCCCGCTCTTTCTTCTTCTCTCTCTCTCTCTCTCTCTCT 1620

Db 1561 GCSCGTCATCTCGTCGTAACGCCCGCTCTTTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCT 1620

Qy 1621 TCGGTGTTTGGTTTCTGTTGGTGGATCCGCTTCCGCGCGCGCTGCTGCCCTGGTCGG 1680

Db 1621 TCGGTGTTTGGTTTCTGTTGGTGGATCCGCTTCCGCGCGCGCTGCTGCCCTGGTCGG 1680

Qy 1681 GAGATGGGAAAGATTGCTGCTGGCTTTGCTGGGATGTTGCGTTAGATCCGTGCAAGCTG 1740

Db 1681 GAGATGGGAAAGATTGCTGCTGGCTTTGCTGGGATGTTGCGTTAGATCCGTGCAAGCTG 1740

Qy 1741 TCTCCCTTTCTTTTGGGTCAAAATCTGGGTTCTTCTTCCGCTCTCTCTCTCTCTCTCTCT 1800

Db 1741 TCTCCCTTTCTTTTGGGTCAAAATCTGGGTTCTTCTTCCGCTCTCTCTCTCTCTCTCTCTCT 1800

Qy 1801 TGCAAAACAAAAATGTTTGGCTTCGAGGGGAAATCTTGTGTAATAATTTACTTGTGTTTCCACC 1860

Db 1801 TGCACACAAAAATGTTGGCTTCGAGGAAATCTTGTAATAATTTACTTGTGTTCCACC 1860  
Qy 1861 TTTTGGAACTGGCGCAATTTTTCATGGTCAAAATTTGCGTTTGTCTTGCACACTTGCAGATT 1920  
Db 1861 TTTTGGAACTGGCGCAATTTTTCATGGTCAAAATTTGCGTTTGTCTTGCACACTTGCAGATT 1920  
Qy 1921 GAGTCGATTTGGGGTTTTTCCATTTGACGGAAGGATATTAGTGACAGAAATCAAGCGAAAG 1980  
Db 1921 GAGTCGATTTGGGGTTTTTCCATTTGACGGAAGGATATTAGTGACAGAAATCAAGCGAAAG 1980  
Qy 1981 AAATATTTATATGATCGGTGTCGACGCAATATAGGATTTTCAAGCGTTTAAAGGTATAGA 2040  
Db 1981 AAATATTTATATGATCGGTGTCGACGCAATATAGGATTTTCAAGCGTTTAAAGGTATAGA 2040  
Qy 2041 ATTTATTTTCGATTTCAAGTAGTATGTTGAAATTTGGAGGTTGCATTTTGTCTCACACA 2100  
Db 2041 ATTTATTTTCGATTTCAAGTAGTATGTTGAAATTTGGAGGTTGCATTTTGTCTCACACA 2100  
Qy 2101 TGTATTTTAAATTTGTCATGTCGCAATGGTGTGAGGATGAGCAAGAAACACAGGC 2160  
Db 2101 TGTATTTTAAATTTGTCATGTCGCAATGGTGTGAGGATGAGCAAGAAACACAGGC 2160  
Qy 2161 GGTTCGTGGTTTTTGTCTGATGCGAAACTTTTAAATAACCAAAACAGAAAGGTGCA 2220  
Db 2161 GGTTCGTGGTTTTTGTCTGATGCGAACTTTTAAATAACCAAAACAGAAAGGTGCA 2220  
Qy 2221 ACTTTATATGATGCAATAAGGTTTGGAAATGTCATCGGAATAAACAAGAGATAGA 2280  
Db 2221 ACTTTATATGATGCAATAAGGTTTGGAAATGTCATCGGAATAAACAAGAGATAGA 2280  
Qy 2281 CGTATTAATCTCTGTTTCAATATTCACCTAAACCAAACTGATCTGTATGGTTATGTA 2340  
Db 2281 CGTATTAATCTCTGTTTCAATATTCACCTAAACCAAACTGATCTGTATGGTTATGTA 2340  
Qy 2341 AACTGTGTGACACAAATTTCAATTAATCTGTTTCTGTTGCGCAATGCTCTGCTTCAG 2400  
Db 2341 AACTGTGTGACACAAATTTCAATTAATCTGTTTCTGTTGCGCAATGCTCTGCTTCAG 2400  
Qy 2401 CAATGCACTACTCTAAATTTTACTGGATCATCAACCAAACTATATGGAAGGTTTATCTA 2460  
Db 2401 CAATGCACTACTCTAAATTTTACTGGATCATCAACCAAACTATATGGAAGGTTTATCTA 2460  
Qy 2461 TTGTTTCTTTTACGTTCAAGAGATATAGCTAAATTTTCAAACTCTCTGCACTACCTAATTT 2520  
Db 2461 TTGTTTCTTTTACGTTCAAGAGATATAGCTAAATTTTCAAACTCTCTGCACTACCTAATTT 2520  
Qy 2521 TAGTGGACCGAATCTAGTCACTGCTGCGCTTAGCTATCTCCAAATTTGATGGCGTTTCT 2580  
Db 2521 TAGTGGACCGAATCTAGTCACTGCTGCGCTTAGCTATCTCCAAATTTGATGGCGTTTCT 2580  
Qy 2581 AATTCATGCAATGCTCACTGAGAACTGGTGTGTTTAAAGCAATTTGTAAGAAATGCCAGTAC 2640  
Db 2581 AATTCATGCAATGCTCACTGAGAACTGGTGTGTTTAAAGCAATTTGTAAGAAATGCCAGTAC 2640  
Qy 2641 ACCCTATATCTTATTCAGAGATAGGATGAGCAGTGAACATAGACGAAGTGTATGCAA 2700  
Db 2641 ACCCTATATCTTATTCAGAGATAGGATGAGCAGTGAACATAGACGAAGTGTATGCAA 2700  
Qy 2701 TTGATGCTTTTATGATGTTATAGCATCTGCCAGAAGCTTGAATATTTGATGCTAGAAATGG 2760  
Db 2701 TTGATGCTTTTATGATGTTATAGCATCTGCCAGAAGCTTGAATATTTGATGCTAGAAATGG 2760  
Qy 2761 TAAGCAATCTCTGGGTCTTAAAAAATAAAGGTAAAGATACAAATTTTGCATCCC 2820  
Db 2761 TAAGCAATCTCTGGGTCTTAAAAAATAAAGGTAAAGATACAAATTTTGCATCCC 2820  
Qy 2821 TTTATTTTACTAAATTTTAGTGTCCACTGCTGCGGAAAGGGAACACAGTCAACCGCTGAT 2880  
Db 2821 TTTATTTTACTAAATTTTAGTGTCCACTGCTGCGGAAAGGGAACACAGTCAACCGCTGAT 2880  
Qy 2881 TAAGATGAATTTTGTCTGTCGCAATTTAGCCACTGGTGATATGTTGAGGCTGCAAGTGC 2940

Db 2881 TAAGATGAATTTTGTCTGTCGCAATTTAGCCACTGGTGATATGTTGAGGGCTGCGAGTGC 2940  
Qy 2941 TGTCTAAACTCCACTTGGGATTAAGGCTAAAGACTATGGACAAGGTAGTTTTTAAAGAA 3000  
Db 2941 TGTCTAAACTCCACTTGGGATTAAGGCTAAAGACTATGGACAAGGTAGTTTTTAAAGAA 3000  
Qy 3001 ACATATAGCAACAGAAATTTATTAACAGAGAGGAATGGGTTTCTTGATCTTTTGTGTTCTTT 3060  
Db 3001 ACATATAGCAACAGAAATTTATTAACAGAGAGGAATGGGTTTCTTGATCTTTTGTGTTCTTT 3060  
Qy 3061 CTTTATCTTCTAGGAGAGCTTGTCTGATGACTTGGTTGTTGGGATTAATGATGAAGC 3120  
Db 3061 CTTTATCTTCTAGGAGAGCTTGTCTGATGACTTGGTTGTTGGGATTAATGATGAAGC 3120  
Qy 3121 CATGAAGAAACTTTCATGCGCAGAAAGGTTTATCTTGATGGTTTCCCTAGAACTGTTGT 3180  
Db 3121 CATGAAGAAACTTTCATGCGCAGAAAGGTTTATCTTGATGGTTTCCCTAGAACTGTTGT 3180  
Qy 3181 TCAAGCACAGAAAGTGAAGTCTTGGTCAATATGACCGCTATATAAAGAGAGCTCTCTTTT 3240  
Db 3181 TCAAGCACAGAAAGTGAAGTCTTGGTCAATATGACCGCTATATAAAGAGAGCTCTCTTTT 3240  
Qy 3241 TGTATTTAGAGCTGCTATATAAATGGAAGTCTTCTATCAITTTGATCACTTTTCTTACTA 3300  
Db 3241 TGTATTTAGAGCTGCTATATAAATGGAAGTCTTCTATCAITTTGATCACTTTTCTTACTA 3300  
Qy 3301 AAAAATGTCAGCTGATGAAATGTTGGCCAAACAAAGTACTAAGATTGACAGGTTCT 3360  
Db 3301 AAAAATGTCAGCTGATGAAATGTTGGCCAAACAAAGTACTAAGATTGACAGGTTCT 3360  
Qy 3361 AAAATTTGCAATTTGATGATCAATATCTGGAAGAAAGAAATTAACCGGTCTGTTGGATCCACC 3420  
Db 3361 AAAATTTGCAATTTGATGATCAATATCTGGAAGAAAGAAATTAACCGGTCTGTTGGATCCACC 3420  
Qy 3421 ATCAAGTGGTAGATCTTATCATCAAAATTTGCTCTCTCTTAAGACTCCTGGAAGTTGATGA 3480  
Db 3421 ATCAAGTGGTAGATCTTATCATCAAAATTTGCTCTCTCTTAAGACTCCTGGAAGTTGATGA 3480  
Qy 3481 TGTAAAGTACATACCAGATTAATCTGCTCTGCTGCTGATCTGCTAGACTCAGATTTTTTAC 3540  
Db 3481 TGTAAAGTACATACCAGATTAATCTGCTCTGCTGCTGATCTGCTAGACTCAGATTTTTTAC 3540  
Qy 3541 CATTTTTCATTTATTTCTAATAGATTTGGTACATATAATGTTGATGCTTTGAGCATATGC 3600  
Db 3541 CATTTTTCATTTATTTCTAATAGATTTGGTACATATAATGTTGATGCTTTGAGCATATGC 3600  
Qy 3601 GCCTTTTACATAATTTCCCTGCTCTCATCATCGATCATCGTAGCTCTGCTTACTTGT 3660  
Db 3601 GCCTTTTACATAATTTCCCTGCTCTCATCATCGATCATCGTAGCTCTGCTTACTTGT 3660  
Qy 3661 TTTAAGACAAACATGAGCCATCTTTTCAAAATTAAGGTTTCAATGAGTACCT 3720  
Db 3661 TTTAAGACAAACATGAGCCATCTTTTCAAAATTAAGGTTTCAATGAGTACCT 3720  
Qy 3721 TCCTCTTGCCACTGTTTAAATTTAGTTCCTTCCAAATAGTAAATAGSTAGTGTGATTAT 3780  
Db 3721 TCCTCTTGCCACTGTTTAAATTTAGTTCCTTCCAAATAGTAAATAGSTAGTGTGATTAT 3780  
Qy 3781 GGATATATTTTCTGTTTGGTTGTTTCTGTCGACGAAGTAAATGCACTTTTGTG 3840  
Db 3781 GGATATATTTTCTGTTTGGTTGTTTCTGTCGACGAAGTAAATGCACTTTTGTG 3840  
Qy 3841 ACAAGAAATGATAGGCAAGTGTGTCATGACACACTTTTCTGCAATTAATG 3900  
Db 3841 ACAAGAAATGATAGGCAAGTGTGTCATGACACACTTTTCTGCAATTAATG 3900  
Qy 3901 CCTGTTTAAAGTGGCCATTAATAGATACATCGACATGTTTTTGTAGCAGAGATTTGATA 3960  
Db 3901 CCTGTTTAAAGTGGCCATTAATAGATACATCGACATGTTTTTGTAGCAGAGATTTGATA 3960  
Qy 3961 TTGTTTCTTATTTGCTTCCAAATTTAAAGCATATCTTCTTTAGCAATGATTTTCAITGGGAC 4020  
Db 3961 TTGTTTCTTATTTGCTTCCAAATTTAAAGCATATCTTCTTTAGCAATGATTTTCAITGGGAC 4020

QY 4021 ATATTGTGCTGCTATTAGTAAATTTGTTGATATCATATATATCTTTTAATTTGGTAAT 4080  
DB |||||||  
QY 4021 ATATTGTGCTGCTATTAGTAAATTTGTTGATATCATATATATCTTTTAATTTGGTAAT 4080  
DB |||||||  
QY 4081 ATTATGTGCACTTCGCTCCCTGATGTTGTTGTTCTTTTCAAAAGGTTACTGAGAAACC 4140  
DB |||||||  
QY 4081 ATTATGTGCACTTCGCTCCCTGATGTTGTTGTTCTTTTCAAAAGGTTACTGAGAAACC 4140  
DB |||||||  
QY 4141 TTAATTTCAAAGGAAGATGACACAGCTGCAGTATTGAAGTCAAGGCTTGAAGCTTCCAC 4200  
DB |||||||  
QY 4141 TTAATTTCAAAGGAAGATGACACAGCTGCAGTATTGAAGTCAAGGCTTGAAGCTTCCAC 4200  
DB |||||||  
QY 4201 GPACAACTAAGCCGTGTAATGTTTCTTTAGCAACTACGTTTTTAAATPATTCAGATATCT 4260  
DB |||||||  
QY 4201 GPACAACTAAGCCGTGTAATGTTTCTTTAGCAACTACGTTTTTAAATPATTCAGATATCT 4260  
DB |||||||  
QY 4261 TTTAGGATGTAGTCTACTTTCAGTTAAGGCGGATTCTTCAGTTGCAATTACAGTGTTC 4320  
DB |||||||  
QY 4261 TTTAGGATGTAGTCTACTTTCAGTTAAGGCGGATTCTTCAGTTGCAATTACAGTGTTC 4320  
DB |||||||  
QY 4321 TGTATATCTTTCATGTTTTT 4341  
DB |||||||  
QY 4321 TGTATATCTTTCATGTTTTT 4341  
DB |||||||

RESULT 2  
AAA10960  
ID AAA10960 standard; DNA; 4341 BP.  
XX  
AC AAA10960;  
DT 14-JUL-2000 (first entry)  
XX  
XX Rice adenylate kinase gene promoter fragment.  
DE  
KW Rice; adenylate kinase; promoter; gene expression control; antibacterial;  
KW insecticide; light resistance; sucrose induction; flower promotion;  
KW increase crop yield; herbicide resistance; ds.  
XX  
OS Oryza sativa.  
XX  
PN WO200015811-A1.  
XX  
XX 23-MAR-2000.  
PD  
PF 10-SEP-1998; 98WO-JP04088.  
XX  
PR 10-SEP-1998; 98WO-JP04088.  
XX  
PA (NISC ) NISSAN CHEM IND LTD.  
XX  
XX Uchimiya H, Arai S, Fushimi T, Tagawa M;  
PI  
XX WPI; 2000-271447/23.  
DR  
XX Rice adenylate kinase gene-originated DNA fragment with promoter  
PT function in plants for controlling expression of desired foreign  
PT structural genes, useful for herbicide resistance and in sucrose  
PT induction in recombinant plants -  
XX  
PS Claim 1; Page 23-25; 30pp; Japanese.  
XX  
CC This sequence represents a rice adenylate kinase gene-originated DNA  
CC fragment with promoter function in plants. The fragment can be used for  
CC controlling the expression of desired foreign structural genes. The DNA  
CC fragment is used to create a vector, which in turn is used to transform  
CC bacteria or plant cells. The DNA fragment can be inserted into a plant to  
CC control the expression of desired foreign structural genes, useful for  
CC inducing in the plant e.g. herbicide, stress and light resistance,  
CC production of antibacterial and insecticidal substances, sucrose  
CC induction in recombinant plants, and in promoting flowering and crop  
CC yield.

XX  
SQ Sequence 4341 BP; 1284 A; 758 C; 880 G; 1419 T; 0 other;  
Query Match 100.0%; Score 4341; DB 21; Length 4341;  
Best local similarity 100.0%; Pred. No. 0;  
Matches 4341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTGCAGGAAGATTAAATTTAGGTGGACACACAAACCCTGTGGTTGGTGGCCCTGTGTT 60  
DB |||||||  
QY 1 CTGCAGGAAGATTAAATTTAGGTGGACACACAAACCCTGTGGTTGGTGGCCCTGTGTT 60  
DB |||||||  
QY 61 AATCAACTGGGTGTTGCTTGGACATGGTTTTGCGAGGAAATTAAGCAGAAATTAAG 120  
DB |||||||  
QY 61 AATCAACTGGGTGTTGCTTGGACATGGTTTTGCGAGGAAATTAAGCAGAAATTAAG 120  
DB |||||||  
QY 121 AAGAATCTCAAGCTGACATGAGAAACCGTAATCCAAATGGAAGCGAATTTCAAGTCGTT 180  
DB |||||||  
QY 121 AAGAATCTCAAGCTGACATGAGAAACCGTAATCCAAATGGAAGCGAATTTCAAGTCGTT 180  
DB |||||||  
QY 181 TCTTGTAATCAATGTTTGAATACATAGACAGTGCCCAACGTTTGAAGCTTCTATTGG 240  
DB |||||||  
QY 181 TCTTGTAATCAATGTTTGAATACATAGACAGTGCCCAACGTTTGAAGCTTCTATTGG 240  
DB |||||||  
QY 241 CTCGTGTGATCTGACTTGTGCACAAAGCATCAAAATTCCTTCTTGGAGTATCTTTATTA 300  
DB |||||||  
QY 241 CTCGTGTGATCTGACTTGTGCACAAAGCATCAAAATTCCTTCTTGGAGTATCTTTATTA 300  
DB |||||||  
QY 301 CCGAAACCCCAAGATATTCTTATTCACCTCAGGTTAAATTTGCTGNACTATGCAATG 360  
DB |||||||  
QY 301 CCGAAACCCCAAGATATTCTTATTCACCTCAGGTTAAATTTGCTGNACTATGCAATG 360  
DB |||||||  
QY 361 AATACAAATTCGCAAAATATCATGTTTATCTATCTTCTGCTCAAAATTTGAGTCCAA 420  
DB |||||||  
QY 361 AATACAAATTCGCAAAATATCATGTTTATCTATCTTCTGCTCAAAATTTGAGTCCAA 420  
DB |||||||  
QY 421 CTGAGACTGCAATACGATTTTCTTTTCAAAAGAAATTTAATTTTTCATGAAA 480  
DB |||||||  
QY 421 CTGAGACTGCAATACGATTTTCTTTTCAAAAGAAATTTAATTTTTCATGAAA 480  
DB |||||||  
QY 481 CGCAATTCAAACCGTTCGAGAAATATCTGTCATAATAGTAGTCTAGTCGAGAAACAAA 540  
DB |||||||  
QY 481 CGCAATTCAAACCGTTCGAGAAATATCTGTCATAATAGTAGTCTAGTCGAGAAACAAA 540  
DB |||||||  
QY 541 ATTAATATCATATAAAAGAGGTTGTTAATTCACAAACCATGTTTCGTACTACAACTCT 600  
DB |||||||  
QY 541 ATTAATATCATATAAAAGAGGTTGTTAATTCACAAACCATGTTTCGTACTACAACTCT 600  
DB |||||||  
QY 601 AATTTGTAAATTCCTTATTTTCAGTCACAAATTCCTTCAATTTTCAAAATTAAGCT 660  
DB |||||||  
QY 601 AATTTGTAAATTCCTTATTTTCAGTCACAAATTCCTTCAATTTTCAAAATTAAGCT 660  
DB |||||||  
QY 661 AGACGGCTAAGCCCATCTTAAGGCTAAGTTTCGAGAGTGAAGTACGACGAAAAATA 720  
DB |||||||  
QY 661 AGACGGCTAAGCCCATCTTAAGGCTAAGTTTCGAGAGTGAAGTACGACGAAAAATA 720  
DB |||||||  
QY 721 TGATGGTTTATTAATATGATTTTAAATAACTTTTCAATATAATTTCTTTAGGAAACA 780  
DB |||||||  
QY 721 TGATGGTTTATTAATATGATTTTAAATAACTTTTCAATATAATTTCTTTAGGAAACA 780  
DB |||||||  
QY 781 TATCATTTAATGTTTGAAGAACGTCACATAGAAACCTAAGAACGATGAGTTCGGAAA 840  
DB |||||||  
QY 781 TATCATTTAATGTTTGAAGAACGTCACATAGAAACCTAAGAACGATGAGTTCGGAAA 840  
DB |||||||  
QY 841 CAAGAGAAAAACACAGCCTTTAGGCTTCTTGATCTCTAGTTGGAGGTTGATTTTCAAAC 900  
DB |||||||  
QY 841 CAAGAGAAAAACACAGCCTTTAGGCTTCTTGATCTCTAGTTGGAGGTTGATTTTCAAAC 900  
DB |||||||  
QY 901 GCATGATAAACGAGAAAGCTCATTAGCACATTTATTACTTAGATATTATTAATAACT 960  
DB |||||||  
QY 901 GCATGATAAACGAGAAAGCTCATTAGCACATTTATTACTTAGATATTATTAATAACT 960  
DB |||||||  
QY 961 TGAATAAATAATTTTATTTGAAATTTTAAACAATGTATGCATAAAATTTTAAAAAAC 1020  
DB |||||||

Db 961 TGAAGAAAAATATTTATTTGAATTTTTTAAACAATGTATGCATAAAATATTTTAAAAAC 1020  
Qy 1021 ACACCAATTTAACCCCTTTAAAAGCATCCTAATAGGAAACGAGGAAGTTAAAGATTCACC 1080  
Db 1021 ACACCAATTTAACCCCTTTAAAAGCATCCTAATAGGAAACGAGGAAGTTAAAGATTCACC 1080  
Qy 1081 GAACTGTTTGGATAATCAAAAAATGGGTGGGATTTAGAAATTTGGTAAATGAATCAGGGTTAG 1140  
Db 1081 GAACTGTTTGGATAATCAAAAAATGGGTGGGATTTAGAAATTTGGTAAATGAATCAGGGTTAG 1140  
Qy 1141 GATTAAATATTAATAATGAAGAGGAGAAATGAATGGTTAGAGTTTAAATGTGCTTTTGG 1200  
Db 1141 GATTAAATATTAATAATGAAGAGGAGAAATGAATGGTTAGAGTTTAAATGTGCTTTTGG 1200  
Qy 1201 GTGGGTAGAAAATATTTTGGCATACACTCCCGAGAGCGGTGGTCTTGGCGTGGCAGA 1260  
Db 1201 GTGGGTAGAAAATATTTTGGCATACACTCCCGAGAGCGGTGGTCTTGGCGTGGCAGA 1260  
Qy 1261 AGCGTCTTTTTTGGTGGAAAAAAAATCTGCTTAAAAAGGAAAAAGAGAGCCAGCTTT 1320  
Db 1261 AGCGTCTTTTTTGGTGGAAAAAAAATCTGCTTAAAAAGGAAAAAGAGAGAGCCAGCTTT 1320  
Qy 1321 GTTGTACCGTCTCACAGAAAACGAAACAAAAAGCCGCCACCTAAACCTCTCTCGATCC 1380  
Db 1321 GTTGTACCGTCTCACAGAAAACGAAACAAAAAGCCGCCACCTAAACCTCTCTCGATCC 1380  
Qy 1381 GACCGAGACTCTCTCAATTTTCAAGCGGCGCACGCGAGAGCAGCGAGTTTCTGTCGAC 1440  
Db 1381 GACCGAGACTCTCTCAATTTTCAAGCGGCGCACGCGAGAGCAGCGAGTTTCTGTCGAC 1440  
Qy 1441 GAACAAGGCTAGTGCAGTAGTTGTGTCGGAGATGGCGGGAACCTGGAGGAGCTGCC 1500  
Db 1441 GAACAAGGCTAGTGCAGTAGTTGTGTCGGAGATGGCGGGAACCTGGAGGAGCTGCC 1500  
Qy 1501 GTCCATCGAGCTGATGACGAGCTGCTCCGCGCATCAAGTCAGCTCAGGCTCAAGCCGACAA 1560  
Db 1501 GTCCATCGAGCTGATGACGAGCTGCTCCGCGCATCAAGTCAGCTCAGGCTCAGGCTCAGAA 1560  
Qy 1561 GCGCGTCATCTCTCGGTCAACGCCGCCCTCTTTCTTCTCTCTCTCTCTCTCTCTCTC 1620  
Db 1561 GCGCGTCATCTCTCGGTCAACGCCGCCCTCTTTCTTCTCTCTCTCTCTCTCTCTCTC 1620  
Qy 1621 TCAGTGGTTTGGTTGGTTGGTGGATCCGGTTTGGCGGCGCGTCTGCGCTGGTGG 1680  
Db 1621 TCAGTGGTTTGGTTGGTTGGTGGATCCGGTTTGGCGGCGCGTCTGCGCTGGTGG 1680  
Qy 1681 GAGATGGGAAAGATTCTCTGTGGCTTCTGCGGATGTTGGTTAGATCCGTGCAAGCTG 1740  
Db 1681 GAGATGGGAAAGATTCTCTGTGGCTTCTGCGGATGTTGGTTAGATCCGTGCAAGCTG 1740  
Qy 1741 TCTCCCTTTCTTTTTTGGGTCAAAATCTGGGTTTCTTCTTCTCTCTCTCTCTCTCTCT 1800  
Db 1741 TCTCCCTTTCTTTTTTGGGTCAAAATCTGGGTTTCTTCTTCTCTCTCTCTCTCTCTCT 1800  
Qy 1801 TGCAAAACAAAATGTTTGGCTTCGAGGGAATCTTGTAAAAATTTACTTGTGTTCACCC 1860  
Db 1801 TGCAAAACAAAATGTTTGGCTTCGAGGGAATCTTGTAAAAATTTACTTGTGTTCACCC 1860  
Qy 1861 TTTGGAACTCGCGCAATTTTGTATGTTCAAAATTTGGGTTTCTTGTGCAACTTGGCAT 1920  
Db 1861 TTTGGAACTCGCGCAATTTTGTATGTTCAAAATTTGGGTTTCTTGTGCAACTTGGCAT 1920  
Qy 1921 GAGTCGATTTGGGTTTTCATTTGACGGAAGGATATAGTGACAGAAATCAAGCGAAAG 1980  
Db 1921 GAGTCGATTTGGGTTTTCATTTGACGGAAGGATATAGTGACAGAAATCAAGCGAAAG 1980  
Qy 1981 AAATATTTATATGATCGGTGTTCCAGCATATAGGAATTTCCAGGCGTTAAGGTATAGA 2040  
Db 1981 AAATATTTATATGATCGGTGTTCCAGCATATAGGAATTTCCAGGCGTTAAGGTATAGA 2040  
Qy 2041 ATTTATTCGATTCAGTAGTATGTTGAAATTTGGAGGTTTGCATTTTTTGTCTCACACA 2100  
Db 2041 ATTTATTCGATTCAGTAGTATGTTGAAATTTGGAGGTTTGCATTTTTTGTCTCACACA 2100

Qy 2101 TGTATATTTTAAATTTGTGCATGTGCAATGGTGTAGGATGAGCAAGAAAAACAAGGCG 2160  
Db 2101 TGTATATTTTAAATTTGTGCATGTGCAATGGTGTAGGATGAGCAAGAAAAACAAGGCG 2160  
Qy 2161 GGTTCGTTGGTTTTTGTCTCATGTGGAACCTTTTAAAAATAACCAAAACAGAAAGTGCA 2220  
Db 2161 GGTTCGTTGGTTTTTGTCTCATGTGGAACCTTTTAAAAATAACCAAAACAGAAAGTGCA 2220  
Qy 2221 ACTTTATATGATTCGATAAAAGTTTGGAAATGTCACCTCGGAATAAAAAACAAGAGATAGA 2280  
Db 2221 ACTTTATATGATTCGATAAAAGTTTGGAAATGTCACCTCGGAATAAAAAACAAGAGATAGA 2280  
Qy 2281 CGTATTACTCTTTTGTTCATATTCACACTACTAAACCAAAATCGATCTGTATGGTTATGTA 2340  
Db 2281 CGTATTACTCTTTTGTTCATATTCACACTACTAAACCAAAATCGATCTGTATGGTTATGTA 2340  
Qy 2341 AACTGTGTGACAAACAAATTTCAAATAAACTTTCTGTTCGCAATGCTCTGCTTCAG 2400  
Db 2341 AACTGTGTGACAAACAAATTTCAAATAAACTTTCTGTTCGCAATGCTCTGCTTCAG 2400  
Qy 2401 CAATGCACTACTCTAAATTTTACTTGGATCATCAAAACAAATCATATGGAAGGTTTATCTA 2460  
Db 2401 CAATGCACTACTCTAAATTTTACTTGGATCATCAAAACAAATCATATGGAAGGTTTATCTA 2460  
Qy 2461 TTCTGTTCTTTTACGTTTCATGAAGTATAGCTTAATTTACAAACTCTCTGCACTACCTAATTT 2520  
Db 2461 TTCTGTTCTTTTACGTTTCATGAAGTATAGCTTAATTTACAAACTCTCTGCACTACCTAATTT 2520  
Qy 2521 TAGTGGACCAATCTAGTCTCTGCGCTTACGCTATCTCCTCAAAATTTGATGGCGTCTCT 2580  
Db 2521 TAGTGGACCAATCTAGTCTCTGCGCTTACGCTATCTCCTCAAAATTTGATGGCGTCTCT 2580  
Qy 2581 AATTCATGCACTGTCACCTGAGAACCTGGTGTGTTTAAACAGCATTTGTAAGATGCCAGTAC 2640  
Db 2581 AATTCATGCACTGTCACCTGAGAACCTGGTGTGTTTAAACAGCATTTGTAAGATGCCAGTAC 2640  
Qy 2641 ACCCTAATATGTTTATGAGGATAGGATAGTGCAGTGTAAACATAGAGCAAGTGTATGCAA 2700  
Db 2641 ACCCTAATATGTTTATGAGGATAGGATAGTGCAGTGTAAACATAGAGCAAGTGTATGCAA 2700  
Qy 2701 TTGATGCTTTAGATAGTTTCTATAGCATCTGCGAAGCTTGAATATTTGATGCTAGAAATGG 2760  
Db 2701 TTGATGCTTTAGATAGTTTCTATAGCATCTGCGAAGCTTGAATATTTGATGCTAGAAATGG 2760  
Qy 2761 TAAGACAATCTCTGCGTGTCTTAAAAAAAAGGTAAAGACATACAAATTTTGACATCCC 2820  
Db 2761 TAAGACAATCTCTGCGTGTCTTAAAAAAAAGGTAAAGACATACAAATTTTGACATCCC 2820  
Qy 2821 TTTATTTTACTAAATTTTAGGTCACCTGGCTAAAGGCTAAGAGCTATGGAAGGAAACACAGTCA 2880  
Db 2821 TTTATTTTACTAAATTTTAGGTCACCTGGCTAAAGGCTAAGAGCTATGGAAGGAAACACAGTCA 2880  
Qy 2881 TAAGGATGAATTTTGTCTGTGCGCATTTAGCCACTGGTGATATGTTGAGGCGTGCAGTGGC 2940  
Db 2881 TAAGGATGAATTTTGTCTGTGCGCATTTAGCCACTGGTGATATGTTGAGGCGTGCAGTGGC 2940  
Qy 2941 TGCTTAAACTCCACTTGGGATTAAGGCTAAGAGCTATGGAAGGAAACACAGTCAAGGAA 3000  
Db 2941 TGCTTAAACTCCACTTGGGATTAAGGCTAAGAGCTATGGAAGGAAACACAGTCAAGGAA 3000  
Qy 3001 ACATATAGCAACAGAAATTTATAACAGGAGGATGGGTTTCTTGATTTCTTTTGTCTTTT 3060  
Db 3001 ACATATAGCAACAGAAATTTATAACAGGAGGATGGGTTTCTTGATTTCTTTTGTCTTTT 3060  
Qy 3061 CTTTATCTTCTAGGAGAGCTTCTTCTGATGCTTGTGTTTGGGATTTATGATGAAC 3120  
Db 3061 CTTTATCTTCTAGGAGAGCTTCTTCTGATGCTTGTGTTTGGGATTTATGATGAAC 3120  
Qy 3121 CATGAAGAAACTTCATGCGAGAAAGGTTTATCTTGTATGGTTTCCCTAGAACTGTTGT 3180  
Db 3121 CATGAAGAAACTTCATGCGAGAAAGGTTTATCTTGTATGGTTTCCCTAGAACTGTTGT 3180



QY 3181 TCAAGCACAGAGGTGAGTCCCTGGTCAATATGCACCGCTATATAAAGAGCTCCTTTT 3240  
Db |||||  
QY 3181 TCAAGCACAGAGGTGAGTCCCTGGTCAATATGCACCGCTATATAAAGAGCTCCTTTT 3240  
Db |||||  
QY 3241 TGTATTAGAGCTCTATATAAATGAGACAGTTTCTTATCATTTGATACACTTTCCTACTA 3300  
Db |||||  
QY 3241 TGTATTAGAGCTCTATATAAATGAGACAGTTTCTTATCATTTGATACACTTTCCTACTA 3300  
Db |||||  
QY 3301 AAAATGGTGCAGCTTGATGAATGTTGGCCAAACAAGGTACTAAGATTGACAAGGTTCCT 3360  
Db |||||  
QY 3301 AAAATGGTGCAGCTTGATGAATGTTGGCCAAACAAGGTACTAAGATTGACAAGGTTCCT 3360  
Db |||||  
QY 3361 AAATTTGGCAATTGATGATCAATACTGGAAGAACAAATACCGGTGCTTGGATCCACCC 3420  
Db |||||  
QY 3361 AAATTTGGCAATTGATGATCAATACTGGAAGAACAAATACCGGTGCTTGGATCCACCC 3420  
Db |||||  
QY 3421 ATCAAGTGTAGATCTTATCATACAAAATTTGCTCTCGTTGCAATTTGTCAGATCTCAGATTTTAC 3480  
Db |||||  
QY 3421 ATCAAGTGTAGATCTTATCATACAAAATTTGCTCTCGTTGCAATTTGTCAGATCTCAGATTTTAC 3480  
Db |||||  
QY 3481 TGTAAAGTATACCAAGATTACTTCTCGTTGCAATTTGTCAGATCTCAGATTTTAC 3540  
Db |||||  
QY 3481 TGTAAAGTATACCAAGATTACTTCTCGTTGCAATTTGTCAGATCTCAGATTTTAC 3540  
Db |||||  
QY 3541 CATTTTTCATTTCTTAATAGATTGGTACATATAAATGGTTGATGCTTGCAGCATATGC 3600  
Db |||||  
QY 3541 CATTTTTCATTTCTTAATAGATTGGTACATATAAATGGTTGATGCTTGCAGCATATGC 3600  
Db |||||  
QY 3601 GCCTTTACCATAATTTCCCTGCTCATCATCGATCATCTGACTCTGCTTACTTGTGTT 3660  
Db |||||  
QY 3601 GCCTTTACCATAATTTCCCTGCTCATCATCGATCATCTGACTCTGCTTACTTGTGTT 3660  
Db |||||  
QY 3661 TTTAAGAACAAACATGAGCCATCACTTTTCAAAAATAAAAAGGTTCAITTGAGTACCT 3720  
Db |||||  
QY 3661 TTTAAGAACAAACATGAGCCATCACTTTTCAAAAATAAAAAGGTTCAITTGAGTACCT 3720  
Db |||||  
QY 3721 TCCTCTGCGACCTGTTAATAGTTAGTTGCTTCCCAATAGTTAAATAGGTAGTGTGATTA 3780  
Db |||||  
QY 3721 TCCTCTGCGACCTGTTAATAGTTAGTTGCTTCCCAATAGTTAAATAGGTAGTGTGATTA 3780  
Db |||||  
QY 3781 GGATATATTTTCTTGTGTTGTTGTTCTCGTACGAAGAGTAAATGACACCTTTGTTG 3840  
Db |||||  
QY 3781 GGATATATTTTCTTGTGTTGTTGTTCTCGTACGAAGAGTAAATGACACCTTTGTTG 3840  
Db |||||  
QY 3841 ACAAGAAATGATAGGACGTTTGGATGACACACATTTTGTGCTTCTGACAAATATG 3900  
Db |||||  
QY 3841 ACAAGAAATGATAGGACGTTTGGATGACACACATTTTGTGCTTCTGACAAATATG 3900  
Db |||||  
QY 3901 CCTGTTTAAGTGGCCATAAATAGATACATCGACATGTTTTTGTAGCAGGAGATTGTATA 3960  
Db |||||  
QY 3901 CCTGTTTAAGTGGCCATAAATAGATACATCGACATGTTTTTGTAGCAGGAGATTGTATA 3960  
Db |||||  
QY 3961 TTGTTTCTATTGCTTCCCAATTAAGATATTTCTTCTTGTAGCAATGATTTTCATGTTGGAC 4020  
Db |||||  
QY 3961 TTGTTTCTATTGCTTCCCAATTAAGATATTTCTTCTTGTAGCAATGATTTTCATGTTGGAC 4020  
Db |||||  
QY 4021 ATATTGCTGCTATTAAGTAAATTTGTTGATATCATATATCTTTTAAATTTGTTAAT 4080  
Db |||||  
QY 4021 ATATTGCTGCTATTAAGTAAATTTGTTGATATCATATATCTTTTAAATTTGTTAAT 4080  
Db |||||  
QY 4081 ATTATGTCACCTTCTGCTCCCTGATTGCTTCTTCTTCAAAAGGTTACTCGAAGAACCC 4140  
Db |||||  
QY 4081 ATTATGTCACCTTCTGCTCCCTGATTGCTTCTTCTTCAAAAGGTTACTCGAAGAACCC 4140  
Db |||||  
QY 4141 TTAATTCAAAAGAAAGATGACACAGCTGCGATTTGAAGTCAAGGCTTTGAAGCCTTCCAC 4200  
Db |||||  
QY 4141 TTAATTCAAAAGAAAGATGACACAGCTGCGATTTGAAGTCAAGGCTTTGAAGCCTTCCAC 4200  
Db |||||  
QY 4201 GTACAACTAGGCTGTATGTTTCTTTTAGCAACTACGTTTTTAAATATTAGATATTCT 4260  
Db |||||  
QY 4201 GTACAACTAGGCTGTATGTTTCTTTTAGCAACTACGTTTTTAAATATTAGATATTCT 4260  
Db |||||  
QY 4261 TTTAGGATGTAGTCTGTTTCAAGTAAAGCGGATTCCTTTCAGTTGCATTTACAGTGTCTCC 4320

Db 4261 TTTAGGATGTAGTCTGTTTCAAGTAAAGCGGATTCCTTTCAGTTGCATTTACAGTGTCTCC 4320  
QY 4321 TGTATATCTTTTCATTTGTTTTT 4341  
Db 4321 TGTATATCTTTTCATTTGTTTTT 4341  
RESULT 3  
AAA07482  
ID AAA07482 standard; DNA; 1440 BP.  
XX  
XX AAA07482;  
XX  
XX 17-JUL-2000 (first entry)  
XX  
DE Rice adenylate kinase DNA sequence fragment #2.  
XX  
KW Rice; adenylate kinase; promoter; structural gene expression;  
KW plant gene expression; herbicide resistance; stress resistance;  
KW light resistance; antibacterial production; flowering induction;  
KW insecticidal substance production; ss.  
XX  
OS Oryza sativa.  
XX  
XX WO200015812-A1.  
XX  
XX 23-MAR-2000.  
XX  
XX 08-SEP-1999; 99WO-JP04847.  
XX  
XX 10-SEP-1998; 98WO-JP04088.  
XX  
XX (NISC ) NISSAN CHEM IND LTD.  
XX  
XX Uchimiya H, Arai S, Fushimi T, Tagawa M, Fukuzawa H;  
XX  
XX WPI; 2000-271448/23.  
XX  
PT Rice adenylate kinase gene-originated DNA fragment with promoter  
PT function in plants for controlling expression of desired foreign  
PT structural genes, useful for herbicide resistance and in sucrose  
PT induction in recombinant plants -  
XX  
PS Claim 10; Page 28; 33pp; Japanese.  
XX  
CC This sequence represents a fragment of the rice adenylate kinase gene.  
CC The invention relates to a DNA fragment (such as this sequence) that has  
CC a regulatory function (preferably acting as a promoter) on the expression  
CC of desired foreign structural genes in a plant when introduced into a  
CC plant body or plant cell. The DNA fragment can be inserted into a plant  
CC to control the expression of desired foreign structural genes, useful for  
CC inducing in the plant e.g. herbicide, stress and light resistance,  
CC production of antibacterial and insecticidal substances, sucrose,  
CC induction of recombinant plants, and in promoting flowering and crop  
CC yield. The DNA fragment has a regulatory function on the expression of  
CC desired structural genes in a plant, particularly with control of the  
CC expression in specific tissues and locations to provide safe chemicals.  
XX  
SQ Sequence 1440 BP; 506 A; 246 C; 272 G; 416 T; 0 other;  
Query Match 33.2%; Score 1440; DB 21; Length 1440;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTGAGGAGAGATTAAATAGGTGGACACACCAACCTGTGGTTGGTGACGCCCTGTGTT 60  
Db |||||  
QY 1 CTGAGGAGAGATTAAATAGGTGGACACACCAACCTGTGGTTGGTGACGCCCTGTGTT 60  
Db |||||  
QY 61 AATCAACTGGGTGTTCTGTGACATGTTTTTGCAGAAAAATTAAAGCAAGAAAAATTAG 120  
Db |||||  
QY 61 AATCAACTGGGTGTTCTGTGACATGTTTTTGCAGAAAAATTAAAGCAAGAAAAATTAG 120  
Db |||||



```
QY 121 AAGATGCTCAAGCTGACATGAGAAAACGTAATCCATCGAAGCGAATTTCAAGTCGTTTC 180
Db 121 AAGATGCTCAAGCTGACATGAGAAAACGTAATCCATCGAAGCGAATTTCAAGTCGTTTC 180
QY 181 TCTTGTAATACTACCTGTTTGAATACATAGACAGTGCCTCAACGTTTGTAGTGCCTCTATTGG 240
Db 181 TCTTGTAATACTACCTGTTTGAATACATAGACAGTGCCTCAACGTTTGTAGTGCCTCTATTGG 240
QY 241 CTGCTGTGATACGCTGTTGTGCACAAAGCATCAAAATGCTTCTTGGAGTATCTTTTATTA 300
Db 241 CTGCTGTGATACGCTGTTGTGCACAAAGCATCAAAATGCTTCTTGGAGTATCTTTTATTA 300
QY 301 CCGAAAAACCCCAAGATTATCTTATTCACCTCAGGCTGAATGCTGCAACTATGCAATG 360
Db 301 CCGAAAAACCCCAAGATTATCTTATTCACCTCAGGCTGAATGCTGCAACTATGCAATG 360
QY 361 AATACAAATTCGCAAAATATCATGTTATCTATCTTGTCTCAAAATGAAATTTGAGTCCAA 420
Db 361 AATACAAATTCGCAAAATATCATGTTATCTATCTTGTCTCAAAATGAAATTTGAGTCCAA 420
QY 421 CTGAGACTGCAATACGATTTTCTTTTCAAAAAGAAATTTATTAATTTTTTTTTCATGAAA 480
Db 421 CTGAGACTGCAATACGATTTTCTTTTCAAAAAGAAATTTATTAATTTTTTTTTCATGAAA 480
QY 481 CGCAATTCACACCGTTTCGAGAAATATGCTGCTATAAATAAGTAGTCTAGTGCAGAAACAAA 540
Db 481 CGCAATTCACACCGTTTCGAGAAATATGCTGCTATAAATAAGTAGTCTAGTGCAGAAACAAA 540
QY 541 ATTAAATATCATATAAAGAGGTTGTTAAATTAACAACCATGTTTCGTACTACAACTCT 600
Db 541 ATTAAATATCATATAAAGAGGTTGTTAAATTAACAACCATGTTTCGTACTACAACTCT 600
QY 601 AATTGTAAATTTCTATTTTCAGTGCACAAATTCCTCAATTTCCAAATTAAGAAAAATAAAGCT 660
Db 601 AATTGTAAATTTCTATTTTCAGTGCACAAATTCCTCAATTTCCAAATTAAGAAAAATAAAGCT 660
QY 661 AGACGGCTAAGCCACCCATCTAAGGCTAAGTTTCGAGAGGTGAAGTACGCACGAAAAATA 720
Db 661 AGACGGCTAAGCCACCCATCTAAGGCTAAGTTTCGAGAGGTGAAGTACGCACGAAAAATA 720
QY 721 TGATGGTTTAAATTAATGATTTTTTTTAAATACTTTTCAATTAATTTCTTTAGGAACA 780
Db 721 TGATGGTTTAAATTAATGATTTTTTTTAAATACTTTTCAATTAATTTCTTTAGGAACA 780
QY 781 TATCATTTAATGTTTGAAAAACGCTGCACATAAGAAAACTTAAGAACGATGAGTTGGAAA 840
Db 781 TATCATTTAATGTTTGAAAAACGCTGCACATAAGAAAACTTAAGAACGATGAGTTGGAAA 840
QY 841 CAAGAGAAAAACACAGCTTAAAGCTTCTTGATCCTCTAGTTGGAGTTGATTTTCAAAC 900
Db 841 CAAGAGAAAAACACAGCTTAAAGCTTCTTGATCCTCTAGTTGGAGTTGATTTTCAAAC 900
QY 901 GCATGATTAACGAGAAAGCTCATTAGCACATTTACTTAGATATTATTAATTAATAACT 960
Db 901 GCATGATTAACGAGAAAGCTCATTAGCACATTTACTTAGATATTATTAATTAATAACT 960
QY 961 TGAATAAATAATTTATTTGAATTTTTTAAACAATGATGATGATGATTTTAAAAAC 1020
Db 961 TGAATAAATAATTTATTTGAATTTTTTAAACAATGATGATGATGATTTTAAAAAC 1020
QY 1021 ACACCAATTTAACCCCTTTAAAGACATCTTAATAGGAACGAGGAGTTAAAGATTCACC 1080
Db 1021 ACACCAATTTAACCCCTTTAAAGACATCTTAATAGGAACGAGGAGTTAAAGATTCACC 1080
QY 1081 GAAGTGTGTAATAATGAAATAATGGGTGGGATTTAGAAATTTGTAATGAATCAGGGTTAG 1140
Db 1081 GAAGTGTGTAATAATGAAATAATGGGTGGGATTTAGAAATTTGTAATGAATCAGGGTTAG 1140
QY 1141 GATTAAATATTAATAAGAGAGGAGGAATGAATGGTTAGAGTTTAAATGTGTCTTTTGG 1200
Db 1141 GATTAAATATTAATAAGAGAGGAGGAATGAATGGTTAGAGTTTAAATGTGTCTTTTGG 1200
QY 1201 GTGGGTAGAAAAATTTATTTGCCATACACTCCCGCGAGCGGTGCGTCTTGGCGTGGCAGA 1260
```

```
Db 1201 GTGGGTAGAAAAATTTATTTGCCATACACTCCCGAGAGCGGTGCGTCTTGGCGTGGCAGA 1260
QY 1261 AGGCTCTTTTTCGTTGGAAAAAACTCTTTTAAAAAGAAAAACAGAAAGAGCCAGCTTT 1320
Db 1261 AGGCTCTTTTTCGTTGGAAAAAACTCTTTTAAAAAGAAAAACAGAAAGAGCCAGCTTT 1320
QY 1321 GTTGTACGCTCTCACAGAAAGGAAACAAAGAGCCGACACCTTAAACCTCTCGATCC 1380
Db 1321 GTTGTACGCTCTCACAGAAAGGAAACAAAGAGCCGACACCTTAAACCTCTCGATCC 1380
QY 1381 GACCGAGACTCTCTCCATTTTCAGCGGCGCACGCGGAGAGCACGCGAGTTCTGTCGAC 1440
Db 1381 GACCGAGACTCTCTCCATTTTCAGCGGCGCACGCGGAGAGCACGCGAGTTCTGTCGAC 1440
```

## RESULT 4

AAA10961

ID AAA10961 standard; DNA; 1440 BP.

XX AAA10961;

XX 14-JUL-2000 (first entry)

XX Partial rice adenylate kinase gene promoter fragment.

XX Rice; adenylate kinase; promoter; gene expression control; antibacterial;  
XX insecticide; light resistance; sucrose induction; flower promotion;  
XX increase crop yield; herbicide resistance; ds.

XX Oryza sativa.

XX WO200015811-A1.

XX 23-MAR-2000.

XX 10-SEP-1998; 98WO-JP04088.

XX 10-SEP-1998; 98WO-JP04088.

XX (NISC) NISSAN CHEM IND LTD.

XX Uchimiya H, Arai S, Fushimi T, Tagawa M;

XX WPI; 2000-271447/23.

XX Rice adenylate kinase gene-originated DNA fragment with promoter  
XX function in plants for controlling expression of desired foreign  
XX structural genes, useful for herbicide resistance and in sucrose  
XX induction in recombinant plants -

XX Example 4; Page 26; 30pp; Japanese.

XX This sequence represents a partial rice adenylate kinase gene-originated  
XX DNA fragment with promoter function in plants. The fragment can be used  
XX for controlling the expression of desired foreign structural genes. The  
XX DNA fragment is used to create a vector, which in turn is used to  
XX transform bacteria or plant cells. The DNA fragment can be inserted into  
XX a plant to control the expression of desired foreign structural genes,  
XX useful for inducing in the plant e.g. herbicide, stress and light  
XX resistance, production of antibacterial and insecticidal substances,  
XX sucrose induction in recombinant plants, and in promoting flowering and  
XX crop yield.

XX Sequence 1440 BP; 506 A; 246 C; 272 G; 416 T; 0 other;

XX Query Match 33.2%; Score 1440; DB 21; Length 1440;

XX Best Local Similarity 100.0%; Pred. No. 0;

XX Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAGGAAGATTAAATTAGGTGGACACACCAACCTGTGGTGGACCCCTGTGTT 60

Db 1 CTGCAGGAAGATTAAATTAGGTGGACACACCAACCTGTGGTGGACCCCTGTGTT 60

```
QY 61 AATCAACTGGGGTTCGTTGGACATGGTTTTGACGAGAAATTAAGCAAGAAATTAAG 120
DB 61 AATCAACTGGGGTTCGTTGGACATGGTTTTGACGAGAAATTAAGCAAGAAATTAAG 120
QY 121 AAGAATGCTCAAGCTGACATGAGAAACGTAATCCAAATGGAAGCAATTTCAAGTCGTTTC 180
DB 121 AAGAATGCTCAAGCTGACATGAGAAACGTAATCCAAATGGAAGCAATTTCAAGTCGTTTC 180
QY 181 TCTTGTAATACCAATGTTTGAATACATACACAGATGCCAAGTTTGATGGCTCCTATTGG 240
DB 181 TCTTGTAATACCAATGTTTGAATACATACACAGATGCCAAGTTTGATGGCTCCTATTGG 240
QY 241 CTCGTGTGATPACTGACTTGTGTACAAAGCATCAAAATGCTTCTTGAGTATCTTTATTA 300
DB 241 CTCGTGTGATPACTGACTTGTGTACAAAGCATCAAAATGCTTCTTGAGTATCTTTATTA 300
QY 301 CCGAAACCCCAAGATTAATCTATCCACCTCAGGGTAATTTGTGCTGAACATGCAATG 360
DB 301 CCGAAACCCCAAGATTAATCTATCCACCTCAGGGTAATTTGTGCTGAACATGCAATG 360
QY 361 AATACAAATTCGCAAAATATCATGGTTATCTATCTGCTCAAAATGAAATTCAGTCCAA 420
DB 361 AATACAAATTCGCAAAATATCATGGTTATCTATCTGCTCAAAATGAAATTCAGTCCAA 420
QY 421 CTGAGACTGCAATACGATTTTCTTTTCAAAAAGAAATTAATTAATTTTTCATGAAA 480
DB 421 CTGAGACTGCAATACGATTTTCTTTTCAAAAAGAAATTAATTAATTTTTCATGAAA 480
QY 481 CGCAATTCACACCGTTCGAGAAATATGCTGCATATAATAGTACTGTAGTCAGAAACAA 540
DB 481 CGCAATTCACACCGTTCGAGAAATATGCTGCATATAATAGTACTGTAGTCAGAAACAA 540
QY 541 ATTAATATCACATAAAAGAAAGTTGTTAATTAACAACCATGTTTCGTACTACAACTCT 600
DB 541 ATTAATATCACATAAAAGAAAGTTGTTAATTAACAACCATGTTTCGTACTACAACTCT 600
QY 601 AATTTGTAATTTCTTATTTTCAGTCACAAAATTCCAAATTTCCAAATTAAGAAATAA 660
DB 601 AATTTGTAATTTCTTATTTTCAGTCACAAAATTCCAAATTTCCAAATTAAGAAATAA 660
QY 661 AGACGGCTAAGCCACCCATCTAAGGCTAAGTTTCGAGAGGTGAAGTACGCACGAAAAATA 720
DB 661 AGACGGCTAAGCCACCCATCTAAGGCTAAGTTTCGAGAGGTGAAGTACGCACGAAAAATA 720
QY 721 TGATGGTTTATTAATATGATTTTTTTTAAATAAATCTTTCACATAAAATTTCTTTAGGAAACA 780
DB 721 TGATGGTTTATTAATATGATTTTTTTTAAATAAATCTTTCACATAAAATTTCTTTAGGAAACA 780
QY 781 TATCATTTTAAATGTTTGAAGAAAGTGCACATAGNAATCTAAGAACGATGAGTTGGGAAA 840
DB 781 TATCATTTTAAATGTTTGAAGAAAGTGCACATAGNAATCTAAGAACGATGAGTTGGGAAA 840
QY 841 CAAGAGAAAAACACAGCCTTTAAGGCTTCTTGATCCCTAGTTGGAGGTTGATTTTCAAAAC 900
DB 841 CAAGAGAAAAACACAGCCTTTAAGGCTTCTTGATCCCTAGTTGGAGGTTGATTTTCAAAAC 900
QY 901 GCATGATAACGAGAAAGCTCATAGCACATTTACTTAGATATTTATTAATTAATAACT 960
DB 901 GCATGATAACGAGAAAGCTCATAGCACATTTACTTAGATATTTATTAATTAATAACT 960
QY 961 TGAAGAAAAATATTTATTTGAATTTTTTAAACATGTATGCATATAATTTTAAAAAAC 1020
DB 961 TGAAGAAAAATATTTATTTGAATTTTTTAAACATGTATGCATATAATTTTAAAAAAC 1020
QY 1021 ACACCAATTTAACCCCTTTTAAAGCATCTTAATAGGAAACGAGGAAAGTTTAAAGATTCACC 1080
DB 1021 ACACCAATTTAACCCCTTTTAAAGCATCTTAATAGGAAACGAGGAAAGTTTAAAGATTCACC 1080
QY 1081 GAAGTCTTTGGGTAATGAAGAAATGGGGTGGGATTTAGAAATGGTAAATGAATCAGGGTTAG 1140
DB 1081 GAAGTCTTTGGGTAATGAAGAAATGGGGTGGGATTTAGAAATGGTAAATGAATCAGGGTTAG 1140
```

```
QY 1141 GATTTAAATATTTAAATGAAAGAGGAGAGAAATGAATGTTAGAGTTTAAATGTTCTTTTGG 1200
DB 1141 GATTTAAATATTTAAATGAAAGAGGAGAGAAATGAATGTTAGAGTTTAAATGTTCTTTTGG 1200
QY 1201 GTGGGTAGAAAAATATTTGCCATACACTCCCGAGAGCGGTGCTTTCGTTGGGCGAGA 1260
DB 1201 GTGGGTAGAAAAATATTTGCCATACACTCCCGAGAGCGGTGCTTTCGTTGGGCGAGA 1260
QY 1261 AGCGTCTTTTTCGTTGGAAAAAACTCTTTAAAAAGGAAACAGAAAAAGCCAGCTTT 1320
DB 1261 AGCGTCTTTTTCGTTGGAAAAAACTCTTTAAAAAGGAAACAGAAAAAGCCAGCTTT 1320
QY 1321 GTTGTACCGCTCTCACAGAAACGAAACAAAAAGCCCAACCTTAAACCTTCTCGATCC 1380
DB 1321 GTTGTACCGCTCTCACAGAAACGAAACAAAAAGCCCAACCTTAAACCTTCTCGATCC 1380
QY 1381 GACCGAGACTCTCCATTTTCAGGGGCGACGCGGAGAGACGCGGAGTTCGTGCGAC 1440
DB 1381 GACCGAGACTCTCCATTTTCAGGGGCGACGCGGAGAGACGCGGAGTTCGTGCGAC 1440

RESULT 5
AAC48619
ID AAC48619 standard; DNA; 809 BP.
XX
AC AAC48619;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 58135.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; terminator; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127482.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
```

PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140881.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-01452913.  
PR 27-JUL-1999; 99US-0145318.  
PR 27-JUL-1999; 99US-0145319.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147935.  
PR 09-AUG-1999; 99US-0148171.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 13-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 03-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 14-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.

```
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match
Best Local Similarity 2.7%; Score 115.8; DB 21; Length 809;
Matches 141; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 3308 GTCAGCTTCATGAAATGTTGGCCAAACAAGGTACTAAGATTGACAAGGTTCTAAATTTT 3367
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 421 GAGAAGCTCGATGAGTGCCTTAAGAGCGGAGGAATGAAATTTGACAAAGTTCTCAACTTT 480
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3368 GCAATTCATGATGCAATCTACTGGAAGAACGAATTACCGGTGCTTGGATCCACCCATCAAGT 3427
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 481 GCTATTGATGCGCAATCTTGGAGGAAGAATAACCGGGGATGGATCCACCCATCGAGT 540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3428 GGTAGATCTTATCATACAAATTTGCTCTCCCTTAAGACTCTGTGACTTGATGATGTAAGT 3487
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 541 GGCAGGAGTTACCACACCAAAATTTGCTCTCCCTCCCAAAACCCCTGGAGTTGATGTGATT 600
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3488 CAT 3490
    ||
DB 601 GAT 603

RESULT 6
AAC48110
ID AAC48110 standard; DNA; 482 BP.
AC AAC48110;
XX
XX
DT 18-OCT-2000 (first entry)
DE
DE Zea mays DNA fragment SEQ ID NO: 56299.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic;
KW pathway; promoter; termination sequence; corn; ss.
XX
XX Zea mays subsp. mays.
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
```

PR	25-OCT-1999;	99US-0161405.	
PR	25-OCT-1999;	99US-0161406.	
PR	26-OCT-1999;	99US-0161359.	
PR	26-OCT-1999;	99US-0161360.	
PR	26-OCT-1999;	99US-0161361.	
PR	28-OCT-1999;	99US-0161920.	
PR	28-OCT-1999;	99US-0161992.	
PR	28-OCT-1999;	99US-0161993.	
PR	29-OCT-1999;	99US-0162142.	
Query Match 2.7%; Score 115.2; DB 21; Length 482;			
Best Local Similarity 80.4%; Pred.No.3e-16;			
Matches 135; Conservative 0; Mismatches 33; Indels 0; Gaps			
QY	2838	TAGGTCACCTGGCTCGGGAAGGGAACACAGTACCGCTGATTAAAGATGAATTTGGCT	2897
Db	271	TCGGCCACCTGGCTCTGGAAAGGGTACTCAGTCTCCCTTTATTAAAGATGAATATTGGC	330
QY	2898	TGTGCCATTTAGCCACTGGTGATATGTTGAGGGCTGCAGTGGCTGCTAAAACTCCACTG	2957
Db	331	TGTGCCATTTAGCCACTGGTGATATGTTGAGGGCTGCTGTGGCAGCCAAGCAACTCTAG	390
QY	2958	GGATTAAAGCTAAAGAAGCTATGGAACAAGTAGTATTTTAAAGAAACATA	3005
Db	391	GTTATCAAGCTAAAGAAGCTATGGAATAAGGGAGAGCTGTTTTCAGATA	438
RESULT 7			
AAC38309			
ID	AAC38309 standard; DNA; 469 BP.		
XX	AAC38309;		
AC			
XX			
DT	17-OCT-2000 (first entry)		
XX			
XX	Zea mays DNA fragment SEQ ID NO: 20515.		
KW	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway; metabolic;		
KW	pathway; promoter; termination sequence; corn; ss.		
OS	Zea mays subsp. mays.		
XX			
XX	EP1033405-A2.		
PN			
PD	06-SEP-2000.		
XX			
PF	25-FEB-2000; 2000EP-0301439.		
XX			
PR	25-FEB-1999;	99US-0121825.	
PR	05-MAR-1999;	99US-0123180.	
PR	09-MAR-1999;	99US-0123548.	
PR	23-MAR-1999;	99US-0125788.	
PR	25-MAR-1999;	99US-0126264.	
PR	29-MAR-1999;	99US-0126785.	
PR	01-APR-1999;	99US-0127462.	
PR	06-APR-1999;	99US-0128234.	
PR	08-APR-1999;	99US-0128714.	
PR	16-APR-1999;	99US-0129845.	
PR	19-APR-1999;	99US-0130077.	
PR	21-APR-1999;	99US-0130449.	
PR	23-APR-1999;	99US-0130510.	
PR	28-APR-1999;	99US-0130891.	
PR	28-APR-1999;	99US-0131449.	
PR	30-APR-1999;	99US-0132048.	
PR	30-APR-1999;	99US-0132407.	
PR	04-MAY-1999;	99US-0132484.	
PR	05-MAY-1999;	99US-0132485.	
PR	06-MAY-1999;	99US-0132486.	
PR	06-MAY-1999;	99US-0132487.	
PR	07-MAY-1999;	99US-0132863.	
PR	11-MAY-1999;	99US-0134256.	
PR	14-MAY-1999;	99US-0134218.	

PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 19-MAY-1999; 99US-0134376.  
PR 20-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139763.  
PR 22-JUN-1999; 99US-0139817.  
PR 23-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140981.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143342.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 23-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.









PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144362.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149922.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.

PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 2.4%; Score 102.6; DB 21; Length 608;  
Best Local Similarity 60.6%; Pred. No. 2.4e-13;

Matches 168; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 2719 CATAGCATCTGCGAAGCTTGAATATTTCATGCTCTAGAAATGTAAGACAATCTCTCGGTG 2778

Db 49 CATGCCAAGCAGCAGTGGCTTCTGTTGATATGGAAGATATTCAGACTGTTGATCTCAT 108

QY 2779 TTCTAAAAAAGGTAAGACATACAAATTTTGACATCCCTTTATTTTACTAAATTTT 2838

Db 109 GTCTGAGCTTCTTCGCCGTATGAAATGTCCCTCTAAACCTGACAAACGCTCGTTTTCAT 168

QY 2839 AGGTCCAGCTCGCGAAGGGAACAGTCACCGCTGATTAAAGGATGAATTTTGCTT 2898

Db 169 CGTCCACCTGTTTCAGGAAGGTACACAGTCTCCGTCTATAAAGGATGAGTTTGT 228

QY 2899 GTGCCATTTAGCCACTGGTGATATGTTGAGGCGCTGCAGTGGCTGTAAACTCCACTTGG 2958

Db 229 GTGTCATTTGCTACTGTGTGACATGCTTTAGAGCTGCTGTGCTGCTAAGACTCCTCTTGG 288

QY 2959 GATTAAGGCTAAAGNAGCTATGGAACAGGTAGTTT 2995

Db 289 TGTGAGGCAAGCAAGCAATGATGAAGGAGAGCTT 325  
RESULT 11  
AAC35687  
ID AAC35687 standard; DNA; 611 BP.  
XX  
AC AAC35687;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 11067.  
XX  
KW Hybridisation assay; Genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.



PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 28-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 08-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144532.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 26-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 26-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 2.4%; Score 102.6; DB 21; Length 1017;  
Best Local Similarity 60.6%; Pred. No. 3e-13;  
Matches 168; Conservative 0; Mismatches 109; Indels 0; Gaps 0;



```
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.

PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.4%; Score 102.6; DB 21; Length 1023;
Best Local Similarity 60.6%; Pred No. 3e-13;
Matches 168; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 2719 CATAGCATCTGGCAGAGCTTGAATATTGATCTCTAGAAATGGAATGGAAGCAATCTTGGGTG 2778
Db 72 CATGGCAACGAGCAGTGGCGCTTCTTGATATGGAAGATATTCAGACCGTGTGATCTCAT 131
QY 2779 TTCTAAAAAAGGTAAGACATACATTTTGACATCCCTTTATTATTACTAAATTT 2838
Db 132 GTCTGAGCTTCTTCGCGGTATGAAATGTGCTCTAAACCTGACAAACGCTCTCGTTTCA 191
QY 2839 AGGTCCACCTGGTGGGAAAGGGAACACAGTCACCGCTGATTAAAGGATGAATTTTGCTT 2898
Db 192 CGGTCCACCTGGTTTCAGGGAAGGTACACAGTCTCCGGTCAATAAAGGATGAGTTTGT 251
QY 2899 GTGCCATTTAGCCACTGGTGATATGTTGAGGCTGACAGTGGCTGTAAACTCCACTTGG 2958
Db 252 GTGTCAATTTGCTTACTGGTGACATGCTTAGAGCTGCTGTGCTGCTAAGACTCCTCTTGG 311
QY 2959 GATTAAGGCTAAAGAGCTATGGACAAGGTAGTTT 2995
Db 312 TGTGAAGGCAAGGAAGCAATGGATAAGGAGAGCTT 348

RESULT 14
ABL72417
ID ABL72417 standard; cdna; 295 BP.
XX ABL72417;
AC ABL72417;
XX
XX
DT 14-MAY-2002 (first entry)
XX
DE Corn tassel-derived polynucleotide (cdps) SEQ ID NO:1791.
XX
KW Corn; corn tassel-derived polynucleotide; cdps; hybrid breeding; CDPs;
KW inheritance; characteristic; growth; development; disease resistance;
KW environmental adaptability; quality; yield; molecular marker;
KW multigene trait; plant breeding; corn tassel; gene; ss.
XX
OS Zea mays.
XX
PN US2001051335-A1.
XX
PD 13-DEC-2001.
XX
PP 16-APR-1999; 99US-0294093.
XX
PR 21-APR-1998; 99US-082567P.
XX
PA (LALG/) LALGUDI R V.
PA (ITOL/) ITO L Y.
PA (SHER/) SHERMAN B K.
XX
PI Lalgudi RV, Ito LY, Sherman BK;
XX
DR WFI; 2002-163647/21.
XX
PT Novel purified corn tassel-derived polynucleotide useful for
PT determining altered gene expression, to recover regulatory elements and
PT to follow inheritance of desirable characteristics through hybrid
PT breeding programs ~
XX
PS Claim 1; SEQ ID 1791; 201pp; English.
XX
CC The present sequence describes a purified corn tassel-derived
CC polynucleotide sequence (cdps) comprising a nucleic acid sequence
```



CC selected from those given in ABL70627 to ABL76833. The cdps sequences  
CC encode corn tassel-derived polypeptides (CDPs). The cdps sequences (I)  
CC can be used for determining altered gene expression, to recover  
CC regulatory elements and to follow inheritance of desirable  
CC characteristics through hybrid breeding programs. (I) are also useful  
CC in the evaluation, and alteration of desired characteristics associated  
CC with growth and development, disease resistance, environmental  
CC adaptability, quality and yield, and as molecular markers for studying  
CC inheritance of multigene traits in a plant breeding program. (I) can be  
CC used to produce a tassel-specific profile of gene transcription, a  
CC transcript image, to clone regulatory elements for use in transformation  
CC vectors, to express a polypeptide, to identify, isolate or extend  
CC identical or related corn tassel nucleic acid sequences from DNA  
CC libraries, in nucleic acid hybridisation or amplification technologies,  
CC as query sequences to determine homology of known sequences, as probe  
CC for use in Southern or Northern hybridisation, and to identify the  
CC presence of and/or to determine the degree of similarity between two  
CC (or more) nucleic acid sequences.

Sequence 295 BP; 69 A; 78 C; 80 G; 68 T; 0 other;

Query Match 2.48; Score 102.2; DB 24; Length 295;

Best Local Similarity 84.1%; Pred. No. 2.2e-13;

Matches 127; Conservative 0; Mismatches 23; Indels 1; Gaps 1;

QY 2838 TAGGTCCACCTGGCTGCGGAAGGGAACACAGTCACCGCTGATTAGGATGAATT-TTGC 2896

[illegible]

**0**

QY 2897 TTGTGCCATTTAGCCACTGGTGATATGTTGAGGGCTGCAGTGGCTGCTAAAACTCCACTT 2956

D**b** 204 CTGTGCCATTAGCCACTGGTGATATGTTAGGGCTGCTGTGCAGCCAAGACAACTCTA 263

---

## RESULT 15

ABN98716/C

ID ABN98716 standard; DNA; 894 BP.

AC ABN98716;

DT 01-AUG-2002 (first entry)

DE Arabidopsis thaliana expressed polynucleotide SEQ ID NO 484.

Arabidopsis thaliana; plant; insecticide; fungicide; transgenic; stress;

KW  
 disease; crop; thale cress; tolerance factor; insect; pathogen; transgene;

KW nutrition; ds.

OS *Arabidopsis thaliana*.

PN US2002023281-A1.

PD 21-FEB-2002.

26-JAN-2001: 2001US-0770445.

27-JAN-2000: 2000US-178472P.

PA (GORL/) GORLACH J.

FA	(GURL/)	GURLA
PA	(ANYYY/)	AN Y:

PA (HAMIL/ ) HAMILTON C M.  
PA (PRIC/ ) PRICE J L.

FA (FRIC/) FRICE J  
PA (RAIN/) RAINES T

PA (YUYY/) YU Y.

PA (RAME//) RAMEA

PA (PAGE/) PAGE A.

PA (MATH/) MATHEW A

PA (LEDF/) LEDFORD B L.

PA (WOES/) WOESSNER J P.

PA	(HAAS/) HAAS W D.	
PA	(GARC/) GARCIA C A.	
PA	(KRIC/) KRICKER M.	
PA	(SLAT/) SLATER T.	
PA	(DAVI/) DAVIS K R.	
PA	(ALLE/) ALLEN K.	
PA	(HOFF/) HOFFMAN N.	
PA	(HURB/) HURBAN P.	
XX		
PI	Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;	
PI	Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;	
PI	Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;	
PI	Hurban P;	

XX  
DR WPT: 2002-400781/43

XX  
PT New Arabidopsis thaliana nucleic acid for identifying homologous

PT New Arabidopsis thaliana nucleic acid for identifying and producing compositions that modulate the expression of nucleic acid sequences

PT producing compositions that modulate the expression of function of its encoded protein, and mapping functional regions of protein -

XX  
PS  
Claim 1. SEQ ID NO 484: 49nn + Sequence Listing. English

PS  
XX  
CLAIM 1; SEQ ID NO 484; 49pb + sequence listing; English.

CC The invention relates to an Arabidopsis thaliana nucleic acid (I)  
CC comprising a sequence capable of hybridising under stringent conditions

comprising a sequence capable of hybridising under stringent conditions to a sequence selected from any one of 999 sequences (ABN98233-ABN99233)

CC given in the specification or its fragment. A polypeptide (II) encoded by (I) a transgenic plant (III) comprising an exogenous nucleic acid or

|||||  
Db 463 CAAAGTGGAGAGAGCTATCATCTAAATTCGACCTCTAAAGTTCGAGAGTCGATGATC 404  
Qy 3483 TAACT 3487  
Db 403 TGAAT 399

Search completed: March 15, 2003, 10:55:20  
Job time : 843.479 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2003, 10:25:55 ; Search time 137.416 Seconds  
(without alignments)  
9687.969 Million cell updates/sec

Title: US-09-802-937-1  
Perfect score: 4341  
Sequence: 1 ctgcaggaagattaattagg.....gtatatcttcattgtttt 4341

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 1533381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	57.2	1.3	7218	1	US-08-232-463-14
C 2	57	1.3	11831	4	Sequence 14, Appl
C 3	53	1.2	7218	1	Sequence 65, Appl
C 4	50.2	1.2	6124	2	US-08-232-463-14
C 5	50	1.2	19124	2	Sequence 14, Appl
C 6	49.2	1.1	669	4	Sequence 13, Appl
C 7	49.2	1.1	19124	2	Sequence 452, Appl
C 8	46.8	1.1	636	4	Sequence 13, Appl
C 9	46.6	1.1	1850	3	Sequence 1137, Ap
C 10	46.6	1.1	2389	1	Sequence 32, Appl
C 11	46.6	1.1	2389	2	Sequence 2, Appl
C 12	46.6	1.1	4098	2	Sequence 4, Appl
C 13	46.6	1.1	4276	1	Sequence 3, Appl
C 14	46.6	1.1	4276	4	Sequence 3, Appl
C 15	46.6	1.1	4276	4	Sequence 3, Appl
C 16	46.6	1.1	4276	4	Sequence 3, Appl
C 17	46.2	1.1	2304	1	US-08-464-266-1
C 18	46.2	1.1	2304	1	Sequence 1, Appl
C 19	46.2	1.1	2304	4	Sequence 1, Appl
C 20	46.2	1.1	2304	4	Sequence 1, Appl
C 21	45.8	1.1	5852	3	US-08-486-403-1
C 22	45.8	1.1	9048	3	US-07-867-106-2
C 23	45.4	1.0	854	2	Sequence 4, Appl
C 24	45.4	1.0	854	2	Sequence 2, Appl
C 25	45.4	1.0	1751	4	US-08-829-027-2
C 26	45	1.0	636	4	Sequence 68, Appl
C 27	44.2	1.0	837	4	Sequence 1137, Ap
					Sequence 288, App

C 28	44	1.0	658	4	US-08-998-416-595
C 29	43.8	1.0	3950	4	US-09-398-395A-33
C 30	43	1.0	1857	4	US-09-299-378-3
C 31	42.8	1.0	6265	4	US-09-129-112-3
C 32	42.6	1.0	20674	4	US-09-641-638-651
C 33	42.6	1.0	28001	4	US-09-819-993-3
C 34	42.4	1.0	665	2	US-08-883-795A-36
C 35	42	1.0	711	4	US-08-998-416-786
C 36	41.8	1.0	3305	3	US-08-714-918-76
C 37	41.8	1.0	3305	4	US-09-265-315-76
C 38	41.8	1.0	3305	4	US-09-265-315-76
C 39	41.6	1.0	1431	4	US-09-316-083-2
C 40	41.6	1.0	2447	2	US-08-973-462-1
C 41	41.6	1.0	6152	4	US-08-973-462-1
C 42	41.4	1.0	665	2	US-08-883-795A-36
C 43	41.4	1.0	2226	4	US-08-924-629C-35
C 44	41.2	0.9	1983	4	US-09-134-001C-2202

ALIGNMENTS

RESULT 1  
US-08-232-463-14/c  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZgpt-F1s  
US-08-232-463-14

Query Match 1.3%; Score 57.2; DB 1; Length 7218;





```
; SEQ ID NO 452
; LENGTH: 669
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-452

Query Match      1.1%; Score 49.2; DB 4; Length 669;
Best Local Similarity 54.4%; Pred. No. 0.0054;
Matches 99; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 3318 ATGAATGTTGCCAAACAAGGTACTAAGATTGACAAGGTTCTAAATTTTGGCAATTGATG 3377
Db      |||||
QY 3311 ATCAAAATTATGCTGAGCTTGATAGAGAAATGATGCTGTCATTATATCGAAGTTCTTG 370
Db      |||||
QY 3378 ATGCAATATCTGAAGAAAGCAATTACCGGTCGTTGATCCACCCATCAAGTGGTAGATCTT 3437
Db      |||||
QY 371 AGGAAGAATTAAATGATCGTCTTACAGGTCGTCGATCTGTGAGAAATGTGGTACAACAT 430
Db      |||||
QY 3438 ATCATACAAAAATTTCTCTCTCTTAAGACTCTCTGAGACTTGTGATGATCTAAGTCTATACCAGAT 3497
Db      |||||
QY 431 ATCACTCTGATTTTAAATCTCTCAAGGTTGATGGTATATGATATCGATCGTGGAAAGT 490
Db      |||||
QY 3498 TA 3499
Db      ||
QY 491 TA 492
Db      ||

RESULT 7
US-08-487-826B-13/c
; Sequence 13, Application US/08487826B
; Patent No. 5993827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,826B
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
```

```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-487-826B-13

Query Match      1.2%; Score 50; DB 2; Length 19124;
Best Local Similarity 45.0%; Pred. No. 0.02;
Matches 273; Conservative 0; Mismatches 325; Indels 8; Gaps 2;

QY 483 CAATTCAACCGTTTCGAGAAATATGCTGTCATAAATAGTAGTCTAGTCAGAGAAACAAAAT 542
Db      |||||
QY 15593 CAAAAAATAATTAATAAATTTTATATAAATAAATAAATGATTATAAATAAATAAATAA 15652
Db      |||||
QY 543 TAAATATACATAAATAAAGGTTGTTAATTAACAACCAATGTTTCTGATCTACAACTCTAA 602
Db      |||||
QY 15653 CAAAAGAAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 15712
Db      |||||
QY 603 TTTGTAAATCTTATTTTCAGTCACAAAATTCCAATTTCCAAATTAAGAAAAATAAACGTA 662
Db      |||||
QY 15713 AAGAATAAATAATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 15771
Db      |||||
QY 663 ACGGCTAAGCCCAACCCATCTAAGGCTTAAGTTCGAGAGGTGAAGTACGCAGCAAAATATG 722
Db      |||||
QY 15772 ATGTTAAAAAATAATATATACATAAATAAATAAATAAATAAATAAATAAATAAATAA 15831
Db      |||||
QY 723 ATGTTTATATATGATTTTTTTTAAATAAATCTTCCATTAATTTCTTTAGGAACATA 782
Db      |||||
QY 15832 TAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 15891
Db      |||||
QY 783 TCATTTAATGTTTGAATAAAGCGTCACATAAGAACTAAGAACGATGATGGGAAACA 842
Db      |||||
QY 15892 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 15951
Db      |||||
QY 843 AGAGAAAAACACAGCCCTTAAGGCTTCTGATCCTCTGATGGAGGTTGATTTTCAACGC 902
Db      |||||
QY 15952 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 16004
Db      |||||
QY 903 ATGATAAAGAGAAAGCTCATTAAGCACATTTACTAGATATTTATATTAATTAACCTTG 962
Db      |||||
QY 16005 ACCATAACTACATAACAACATTTTACACATACATATATATATATATATATATATAT 16064
Db      |||||
QY 963 AAAAAAATAATTTTGAATTTTAAACAATGATGCATAAATAATTTTAAACAAACAC 1022
Db      |||||
QY 16065 CATACATTTTACACATACATATATATATATATATATATATATATATATATATATAT 16124
Db      |||||
QY 1023 ACCAATTTAACCTTTAAAGAGCATCTTAATAGGAAACGAGGAAGTTAAAGATTACCGA 1082
Db      |||||
QY 16125 ATACATATATACATTAAACAACACATATATAATACCTAAATACATATATATATATAT 16184
Db      |||||
QY 1083 AGTGT 1088
Db      |||||
QY 16185 TATGTT 16190
Db      |||||

RESULT 6
US-09-134-001C-452
; Sequence 452, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
```



US-08-487-826B-13

Query Match 1.1%; Score 49.2; DB 2; Length 19124;  
Best Local Similarity 47.3%; Pred. No. 0.032;  
Matches 244; Conservative 0; Mismatches 268; Indels 4; Gaps 3;  
QY 3604 TTATACATATTTCCCTGCTCATCATCGATCATCGTAGTACTCTGCTACTGTTT 3663  
DB 15945 TTTAAATAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 15886  
QY 3664 AAGAACAAACATGAGCCATCTCTTCAAAATAAAGAGTTCATTTGAGTACCTCC 3723  
DB 15885 AATAAATTTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTAT 15826  
QY 3724 TCTGCCACTGTTTAAATAGTTGCTTCCAAATAGTAAATAGGTAAGTATGGA 3783  
DB 15825 TTTTATTTAAATAAATTTTATTTATTTATTTATTTATTTTATTTTAAATTT 15766  
QY 3784 TATATTTTCTGTTGGTGTTCCTCGTACGAAGAGTAAA--ATGCACCTTTTGTGA 3841  
DB 15765 AATTTTATTTTATTTATGATATATATTTTATTTTATTTTATTTTATTTT 15706  
QY 3842 CAAGAATGATAGGCTGTTTGCATGACACACTTTTGTCTTCTTCTGACATATGC 3901  
DB 15705 GTTTTATGATATATATTTTATTTTATTTTATTTTATTTTATTTTCTTT 15646  
QY 3902 CTGTTTAAAGTGGCCATAAATAGATACATGACATGTTTGTGAGGAGATGTATAT 3961  
DB 15645 TTTTATTAATCATTTTATTTTATTTATTTTATTTTATTTTATTTTATTTT 15587  
QY 3962 TGTTTCTATTTCTCCAAATPAAAGCATATCTTCTTTAGCAATGATTCATGGGACA 4021  
DB 15586 TTTTTCATTTTATTTATCTATCAAAATTTATTTTATTTATTTATTTTATTT 15527  
QY 4022 TATTTGCTGCTATTAAGTAAATTTGTTTGATATCATATATATCTTTAAATGGTAATA 4081  
DB 15526 AATTTTCTCCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 15468  
QY 4082 TTATGTCACCTTCGCTCCCTGATGTTGTTGCTTT 4117  
DB 15467 CATTTTCTTTTTCATTTTAAATAAATTTGTTTT 15432

RESULT 8

US-08-998-416-1137/c  
; Sequence 1137; Application US/08998416  
; Patent No. 6239264  
; GENERAL INFORMATION:  
; APPLICANT: Philippsen, Peter  
; APPLICANT: Pohlmann, Rainer  
; APPLICANT: Steiner, Sabine  
; APPLICANT: Mohr, Christine  
; APPLICANT: Wendland, Jurgon  
; APPLICANT: Knechtle, Philipp  
; APPLICANT: Reischung, Corinne  
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPHII  
; TITLE OF INVENTION: AND USES THEREOF  
; NUMBER OF SEQUENCES: 1152  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 6239264artis Corporation  
; STREET: 3054 Cornwallis Road  
; CITY: Research Triangle Park  
; STATE: No. 6239264th Carolina  
; COUNTRY: USA  
; ZIP: 27709  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/998,416  
; FILING DATE: 24-DEC-1997

CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: CH 0016/97  
; FILING DATE: 31-DEC-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meigs, J. Timothy  
; REGISTRATION NUMBER: 38,241  
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 919-541-8587  
; TELEFAX: 919-541-8689  
; INFORMATION FOR SEQ ID NO: 1137:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 636 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ORIGINAL SOURCE:  
; ORGANISM: PAG1692RP  
; US-08-998-416-1137

Query Match 1.1%; Score 46.8; DB 4; Length 636;  
Best Local Similarity 44.6%; Pred. No. 0.021;  
Matches 274; Conservative 0; Mismatches 332; Indels 8; Gaps 2;  
QY 446 TTCAAAAGAAATTAATTTTTCATGAAACGCAATTCACCGTTTCGAGAAATAT 505  
DB 636 TTTTATAGTATTTTAACTACATTCITTATATATTTTATTAATTAATTAATGAT 577  
QY 506 GCTGTCATAAATAGTAGTCTAGTGCAGAAACAAATTAATATCACATAAAAGAGGT 565  
DB 576 AAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 517  
QY 566 TCTTAATTAACAAACCATGTTTCGTACTACACTCTAAATTCGTAAATTCCTTATTCAGTCA 625  
DB 516 TATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 457  
QY 626 CAAAATTCCAATTTCCAAATTAAGAAAAATAAACGTAGACGGCTAAGCCACCCACTCTAAG 685  
DB 456 TCTTATATAAAGATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 397  
QY 686 GCTAAGTTCGAGAGTGAAGTACGCACGAAATAATATGATGGTTTATTA-----ATATG 738  
DB 396 ATAAAAATAAATTTACATAATTTTAAATAAATTTAAATCTTTTATAATAATAATAATAAT 337  
QY 739 ATTTTATTTTAAATAACTTTTCACATAAATTTCTTTAGGAAACATATCATTTTAATGGTTGA 798  
DB 336 ATTTTAAATAACAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 277  
QY 799 AAAACGTGCACATRAAGAAACCTAAGAACGATGAGTTGGGAAACAGAGAGAAACACAGCC 858  
DB 276 TAAAGAAAATAATAATCTAATAATAATTTTAAATACTAATTTTAAATTTTGAACATAGAC 217  
QY 859 TTAAGGCTTCTTGATCCCTCTAGTTGGAGGTTGATTTTCAAACGCGATGATAAACGAGAAAG 918  
DB 216 TAAATAGTATTCATATTAATAATAATAATAATAATAATAATAATAATAATAATAATGAT 157  
QY 919 CTCATTAGCACATTTACTTTAGTATTTTATAATTTATAAATCTGAAAAAATAATTTATTT 978  
DB 156 TAAATGATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 98  
QY 979 GAATTTTAAACAAATGATGCAATAAATTTTATTTTAAACCAACACACCAATTTAAACCTTT 1038  
DB 97 AATAATGATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 38

RESULT 9

US-08-617-860B-32/c



Db 2075 AACATGAATACCTGATCGCACACAGAGCATTGAGAGCAGAGAAAGAAATCGATTGAGA 2016  
Qy 899 ACCCATGATAAGAGAAAGCTCATTAGCACATTTACTTAGATATTTATTAATATAAA 958  
Db 2015 ATGTAGCGAAACGAGCTGAGCCATTAGCTATAATTAATGAGTATTAATCTTTTAA 1956  
Qy 959 CTTGAAAAAATATTTATTTGAATTTTTTAAACAATGATGCATATAATTTTAAAA 1018  
Db 1955 TTTTAAAAATAGATTAATAATTTTAAAGTAACCTTCTCTATAGAAAATTTTGCAAAA 1896  
Qy 1019 ACACACCAATTTAACCCCTTTAAAA 1043  
Db 1895 ATCATACCGTTTAGTAGTTCAGGAA 1871

RESULT 11  
US-08-460-507-2/c  
; Sequence 2, Application US/08460507  
; Patent No. 5984628  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: PROCESS FOR PROTEIN PRODUCTION  
; TITLE OF INVENTION: IN PLANTS  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: 350 Cambridge Avenue, Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/460.507  
; FILING DATE:  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/153,563  
; FILING DATE: 16-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fabian, Gary R.  
; REGISTRATION/DOCKET NUMBER: 33,875  
; REFERENCE/DOCKET NUMBER: 2000-0452.41  
; TELEPHONE: (415) 324-0880  
; TELEFAX: (415) 324-0960  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2389 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..2389  
; OTHER INFORMATION: /standard\_name="RAMV-1A"

Query Match 1.1%; Score 46.6; DB 2; Length 2389;  
Best Local Similarity 51.7%; Pred. No. 0.049;  
Matches 106; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

Qy 839 AACAAAGAAAAACACAGCCTTAAGCCTTCTTGATCCTCTAGTTGGAGGTTGATTTCAA 898  
Db 2075 AACATGAATACCTGATCGCACACAGAGCATTGAGAGCAGAGAAAGAAATCGATTGAGA 2016  
Qy 899 ACCCATGATAAGAGAAAGCTCATTAGCACATTTACTTAGATATTTATTAATATAAA 958

Db 2015 ATGTACCAAAACGAGCTGAGCCATTAACTAATTAATTGAGTATTAACTATTTTAA 1956  
Qy 959 CTTGAAAAAATATTTATTTGAATTTTTTAAACAATGATGCATATAATTTTAAAA 1018  
Db 1955 TTTTAAAAATAGATTAATAATTTTAAAGTAACCTTCTCTATAGAAAATTTTGCAAAA 1896  
Qy 1019 ACACACCAATTTAACCCCTTTAAAA 1043  
Db 1895 ATCATACCGTTTAGTAGTTCAGGAA 1871

RESULT 12  
US-08-605-106-4/c  
; Sequence 4, Application US/08605106  
; Patent No. 5910631  
; GENERAL INFORMATION:  
; APPLICANT: Topfer, R.  
; APPLICANT: Martini, N.  
; APPLICANT: Schell, J.  
; TITLE OF INVENTION: MEDIUM CHAIN-SPECIFIC THIOESTERS  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.  
; STREET: P.O. Box 2938  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/605,106  
; FILING DATE: 23-SEPT-1996  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/EP94/02935  
; FILING DATE: 01-MAR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Woessner, Warren D  
; REGISTRATION NUMBER: 30,440  
; REFERENCE/DOCKET NUMBER: 235.001US1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612-373-6900  
; TELEFAX: 612-339-3061  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4098 Base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double stranded  
; TOPOLOGY: linear  
; MOLECULE TYPE: : DNS (genomic)  
; HYPOTHETICAL: No  
; ANTI-SENSE: No  
; ORIGINAL SOURCE:  
; ORGANISM: Cuphea lanceolata  
; IMMEDIATE SOURCE:  
; LIBRARY: genomic Lambda FIX II  
; CLONE: C1Eg1  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: Join(1797..2294, 2658..2791, 2898..3011, 3132  
; LOCATION: ..3303, 3391..3459, 3672..3941)  
; FEATURE:  
; NAME/KEY: Startcodon  
; LOCATION: 1797..1799  
; FEATURE:  
; NAME/KEY: exon II  
; LOCATION: 1787..2294



Qy 1019 ACACACCAATTTAACCCCTTTAAAAA 1043  
Db 1889 ATCATACCGTTTAGTAGTTCAGAA 1865

## RESULT 14

US-08-343-380-3/c  
; Sequence 3, Application US/08343380  
; Patent No. 5712112  
; GENERAL INFORMATION:  
; APPLICANT: Yu, Su-May  
; APPLICANT: Liu, Li-Fei  
; TITLE OF INVENTION: Gene Expression System Comprising the  
; TITLE OF INVENTION: Promoter Region of the Alpha-Amylase Genes  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/343.380  
; FILING DATE: 22-NOV-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/973,324  
; FILING DATE: 04-NOV-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Borun, Michael F.  
; REGISTRATION NUMBER: 25447  
; REFERENCE/DOCKET NUMBER: 31149  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4276 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ORIGINAL SOURCE:  
; ORGANISM: Rice (Oryzae sativa)  
; STRAIN: CV. M202  
; IMMEDIATE SOURCE:  
; LIBRARY: (EMBL) genomic  
; CLONE: '-Amy7-C  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: join(2459..2473, 2582..2713, 2807..3619, 3704  
; LOCATION: ..3952)  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: join(2459..2473, 2582..2713, 2807..3619, 3704  
; LOCATION: ..3952)  
; PUBLICATION INFORMATION:  
; AUTHORS: Yu et al., Su-May  
; TITLE: Regulation of '-amylase-encoding gene expression  
; TITLE: in germinating seeds and cultured cells of rice  
; JOURNAL: Gene  
; VOLUME: in press  
US-08-343-380-3

Query Match 1.1%; Score 46.6; DB 1; Length 4276;

Best Local Similarity 51.7%; Pred. No. 0.066;  
Matches 106; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

Qy 839 AACAGAGAGAAAAACACAGCCTTAAGCCTTCTTGATCCTCTAGTTGGAGGTTGATTTCAA 898  
Db 2069 AACATGAATACCTGATCGCACACAGAGCATTTGAGAGAGAGGAAAGATCGATTGAGA 2010  
Qy 899 ACGCATGATAACGAGAAAGCTCATTAGCACATTATTACTTAGATATTTATAATATAAA 958  
Db 2009 ATGTACGCAAAACGAGCTGAGCCATTAACTAATAATTGAGTATTAACTATTTTAAA 1950  
Qy 959 CTTGAAAAAATATTATTGATTTTAAACATGATGCATAAATATTATTTTAAAA 1018  
Db 1949 TTTTAAAAATAGATTATAATATTTTAAAGTAACTTCTCTATAGAAAAATTTTGCAAAA 1890  
Qy 1019 ACACACCAATTTAACCCCTTTAAAAA 1043  
Db 1889 ATCATACCGTTTAGTAGTTCAGAA 1865

## RESULT 15

US-09-072-435-3/c  
; Sequence 3, Application US/09072435  
; Patent No. 6215051  
; GENERAL INFORMATION:  
; APPLICANT: Yu, Su-May  
; APPLICANT: Liu, Li-Fei  
; APPLICANT: Chan, Ming-Tsair  
; TITLE OF INVENTION: GENE EXPRESSION SYSTEM COMPRISING THE  
; TITLE OF INVENTION: PROMOTER REGION OF THE ALPHA-AMYLASE GENES  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive/6300 Sears Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/072,435  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/639,792  
; FILING DATE: 29-APR-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/973,324  
; FILING DATE: 04-NOV-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gass, David A.  
; REGISTRATION NUMBER: 38,153  
; REFERENCE/DOCKET NUMBER: 28123/34274  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4276 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ORIGINAL SOURCE:  
; ORGANISM: Rice (Oryzae sativa)  
; STRAIN: CV. M202  
; IMMEDIATE SOURCE:  
; LIBRARY: (EMBL) genomic  
; CLONE: '-Amy7-C

```
;
;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(2459..2473, 2582..2713, 2807..3619, 3704
; LOCATION: ..3952)
;
; FEATURE:
; NAME/KEY: mat.peptide
; LOCATION: join(2459..2473, 2582..2713, 2807..3619, 3704
; LOCATION: ..3952)
;
; PUBLICATION INFORMATION:
; AUTHORS: Yu et al., Su-May
; TITLE: Regulation of -amylase-encoding gene expression
; TITLE: in germinating seeds and cultured cells of rice
; JOURNAL: Gene
; VOLUME: in press
;
US-09-072-435-3

Query Match      1.1%; Score 46.6; DB 4; Length 4276;
Best Local Similarity 51.7%; Pred. No. 0.066;
Matches 106; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 839 AACAGAGAGAAAACACAGCCTTAAGGCTTCTTGATCCTCTAGTTGGAGGTTGATTTTCAA 898
Db 2069 ACATGAATACCTGATCGCACAGAGCATTTGAGAGCAGAGAAAAGAAATCGATTGAGA 2010

QY 899 ACGCATGATAACGAGAGAAAGCTCATTAGCACATTATTACTTAGATATTTTATAATTATATA 958
Db 2009 ATGTACGCAAAACGAGCTGAGCCATTAAACGTATAATTAATTGAGTATTAACTATTTTAAA 1950

QY 959 CTTGAAATAATATTTATTGAAATTTTAAACAATGATGCATAAATTTTTTTTAAAA 1018
Db 1949 TTTTAAAAATAGATTATATATAATTTTAAAGTAACTTCCTATAGAAAATTTTGCAAAAA 1890

QY 1019 ACACACCAATTTAACCCCTTTAAAAA 1043
Db 1889 ATCATACCGTTTAGTAGTTCAGGAA 1865

Search completed: March 15, 2003, 16:59:03
Job time : 365.416 secs
```

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	3987.8	91.9	4354	10	US-09-802-927-1	Sequence 1, Appli
2	1384.6	31.9	1443	10	US-09-802-927-2	Sequence 2, Appli
3	122	2.8	404	10	US-08-878-574-3892	Sequence 3892, Ap
C	112.8	2.6	981	10	US-09-770-445-271	Sequence 271, App
5	102.2	2.4	295	10	US-09-294-093B-1791	Sequence 1791, Ap
C	101.8	2.3	894	10	US-09-720-445-484	Sequence 484, App
7	92.8	2.1	292	10	US-09-923-876-3999	Sequence 3999, Ap
8	92	2.1	267	10	US-09-923-876-6281	Sequence 6281, Ap
C	62.4	1.4	2000	10	US-09-887-576-828	Sequence 828, App
10	61.6	1.4	645	10	US-08-815-242-6911	Sequence 6911, Ap
11	57	1.3	639	10	US-09-815-242-9121	Sequence 9121, Ap
12	57	1.3	908	12	US-10-044-090-213	Sequence 213, App
C	55.8	1.3	8416	8	US-08-910-386A-4	Sequence 4, Appli
14	55.8	1.3	1341	8	US-08-910-386A-1	Sequence 1, Appli
C	55.8	1.3	19639	8	US-08-910-386A-6	Sequence 6, Appli
16	55.6	1.3	355	10	US-09-860-352-11750	Sequence 11750, A
17	53.6	1.2	398	10	US-09-860-352-11004	Sequence 11004, A
C	53.2	1.2	5940	8	US-08-910-386A-10	Sequence 10, Appli
19	52.2	1.2	238	10	US-09-864-761-31345	Sequence 31345, A

Qy 181 TCTTGTAATACTGTTTGAATACATAGACAGTGCCAAAC-GTTTGTAGTGCTCCTATTG 239  
Db 181 TCTTGTAATACTGTTTGAATACATAGACAGTGCCAAACGTTTGTAGTGCTCCTATTG 240  
Qy 240 GCTCGTGATGACTGCTGTGTACAAAGCATCAAAATGCTTCTTGGAGTATCTTTATT 299  
Db 241 GCTCGTGATGACTGCTGTGTACAAAGCATCAAAATGCTTCTTGGAGTATCTTTATT 300  
Qy 300 ACCGAAACCCCAAGATATTCTATTCCACCTCAGGGTAATGTGCTGAACTATGCAAT 359  
Db 301 ACCGAAACCCCAAGATATTCTATTCCACCTCAGGGTAATGTGCTGAACTATGCAAT 360  
Qy 360 GAATACAAATTCGCAAAATATCATGGTTATCTATCTTGTCTCAAAATGAAATTTGAGTCCA 419  
Db 361 GAATACAAATTCGCAAAATATCATGGTTATCTATCTTGTCTCAAAATGAAATTTGAGTCCA 420  
Qy 420 ACTGAGACTGCAATACGATTTTTCTTTTCAAAAAGAAATATTAAATTTTTTTTTTTCATGAA 479  
Db 421 ACTGAGACTGCAATACGATTTTTCTTTTCAAAAAGAAATATTAAATTTTTTTTTTTCATGAA 480  
Qy 480 ACCTAATTCACCGTTCGAAATATGCTGTCAATAATAGTAGTCTAGTCGAGAAACAA 539  
Db 481 ACCTAATTCACCGTTCGAAATATGCTGTCAATAATAGTAGTCTAGTCGAGAAACAA 540  
Qy 540 AATTAAATACATATAAAAGAGGTGTTAAATTACAAACCATGTTTCGTACTCAAACTC 599  
Db 541 AATTAAATACATATAAAAGAGGTGTTAAATTACAAACCATGTTTCGTACTCAAACTC 600  
Qy 600 TAATTTGTAATTTCTATTTCAGTCAAAATTTCCAAATTTCCAAATTTAGAAATAAAG 659  
Db 601 TAATTTGTAATTTCTATTTCAGTCAAAATTTCCAAATTTCCAAATTTAGAAATAAAG 660  
Qy 660 TAGACGCTAAGCCCAACCATCTAAGCTAAGTTTCGAGAGGTGAAGTACGACGAAAT 719  
Db 661 TAGACGCTAAGCCCAACCATCTAAGCTAAGTTTCGAGAGGTGAAGTACGACGAAAT 720  
Qy 720 ATGATGTTTATTAATATGATTTTTTTTAAATAACTTTTCACTAATAATTTCTTTAGGAAAC 779  
Db 721 ATGATGTTTATTAATATGATTTTTTTTAAATAACTTTTCACTAATAATTTCTTTAGGAAAC 780  
Qy 780 ATATCATTTAATGTTTGAACACGTGCATAGAAACCTAAGACGATGATGTTGGAA 839  
Db 781 ATATCATTTAATGTTTGAACACGTGCATAGAAACCTAAGACGATGATGTTGGAA 838  
Qy 840 ACAAGAGAAACACACGCTTAAAGGCTTCTGTATCTCTAGTTGGAGGTTGATTTCAA 899  
Db 839 ACAAGAGAAACACACGCTTAAAGGCTTCTGTATCTCTAGTTGGAGGTTGATTTCAA 898  
Qy 900 CGCATGATAACGAGAAAGCTCAATAGCACATTTACTTAGATATTTAATATTAAC 959  
Db 899 CGCATGATAACGAGAAAGCTCAATAGCACATTTACTTAGATATTTAATATTAAC 958  
Qy 960 TTGAAAAAATATTATTGCAATTTTTTAAACAAATGATGCATAATTTATTTTTTAAAA 1019  
Db 959 TTGAAAAAATATTATTGCAATTTTTTAAACAAATGATGCATAATTTATTTTTTAAAA 1018  
Qy 1020 CACACCAATTTAACCCTTTTAAAGCATCTTAATAGGAAACGAGGAGTTAAAGATTAC 1079  
Db 1019 CACACCAATTTAACCCTTTTAAAGCATCTTAATAGGAAACGAGGAGTTAAAGATTAC 1078  
Qy 1080 CGAA--GTGTTTGGATAATGAAAAATGGGGTGGGATTAGAATTGGTAATGAATCAGG 1136  
Db 1079 CGAAGTGCTGTTGGATAATGAAAAATGGGGTGGGATTAGAATTGGTAATGAATCAGG 1138  
Qy 1137 TTAGGATTAATAATTTAAATGAAAGGGGAGAAATGAATGGTTAGAGTTTAAATGTGCTT 1196  
Db 1139 TTAGGATTAATAATTTAAATGAAAGGGGAGAAATGAATGGTTAGAGTTTAAATGTGCTT 1198  
Qy 1197 TTTGGGGTAGAAAAATTATTTGCCATACATCCCGAGAGCGGTCGCTTGGTGGG 1256  
Db 1199 TTTGGGGTAGAAAAATTATTTGCCATACATCCCGAGAGCGGTCGCTTGGTGGG 1258  
Qy 1257 CAGAAGCGTCTTTTTCGTTGGAAAAAAACCTGCTTTTAAAAAGGAAACAGAAAGGCCCG 1316

Db 1259 CAGAAGCGTCTTTTTCGTTGGAAAAAAACCTGCTTTAAAAAGGAAACAGAAAGGCCCG 1318  
Qy 1317 CTTT-GTTTGTCACCGTCTCACAGAAACGAAACAAAAAGCCCCACACCTAAACCTCTCTC 1375  
Db 1319 CTTTGGTTGTACCGTCTCACAGAAACGAAACAAAAAGCCCCACACCTAAACCTCTCTC 1378  
Qy 1376 GATCCGACCGAGACTCCTCCATTTTCAGCGCGCAACGCGAGAGCAACGCAACCTGAGGAC 1435  
Db 1379 GATCCGACCGAGACTCCTCCATTTTCAGCGCGCAACGCGAGAGCAACGCGACGAGTTG 1438  
Qy 1436 TCAGACGAAACAAAGCTAGTCAGTAGTTGTTGTCGCGAGATGCGCGCAACCTGAGGAC 1495  
Db 1439 TCAGACGAAACAAAGCTAGTCAGTAGTTGTTGTCGCGAGATGCGCGCAACCTGAGGAC 1498  
Qy 1496 GTCCGCTCGATGAGCTGATGACGAGCTGCTCCGCGCGCATGAAGTGCAGCTCCAAGGCC 1555  
Db 1499 GTCCGCTCGATGAGCTGATGACGAGCTGCTCCGCGCGCATGAAGTGCAGCTCCAAGGCC 1558  
Qy 1556 GACAAGCGCTCATCTCGTCGTAACGCGCGCTCTTCTTCTTCTCTCTCTCTCTCTCTCT 1615  
Db 1559 GACAAGCGCTCATCTCGTCGTAACGCGCGCTCTTCTTCTTCTCTCTCTCTCTCTCTCT 1618  
Qy 1616 CTCTCTCGTGGTTTGGTTTGGTTGGATCCGCTTCGCGCGCGCGCTGCTGCCCCG 1675  
Db 1619 CTCTCTCGTGGTTTGGTTTGGTTGGATCCGCTTCGCGCGCGCGCTGCTGCCCCG 1678  
Qy 1676 GTCCGAGATGCGGAAAGATGCTGTGGCTTCTGCGGAGATTTGGCTTAGATCCGTGCA 1735  
Db 1679 GTCCGAGATGCGGAAAGATGCTGTGGCTTCTGCGGAGATTTGGCTTAGATCCGTGCA 1738  
Qy 1736 AGCTGCTCCCTTCTTTTGGGTCAAACTCTGGGTCTTCTTCTTCCGCTCTCTCTCTGG 1795  
Db 1739 AGCTGCTCCCTTCTTTTGGGTCAAACTCTGGGTCTTCTTCTTCCGCTCTCTCTGG 1798  
Qy 1796 TTTTTCGAAACAAAAATGTTTGGCTTCGAGGAAATCTTGTAAAAATTTTACTTGTGTTT 1855  
Db 1799 TTTTTCGAAACAAAAATGTTTGGCTTCGAGGAAATCTTGTAAAAATTTTACTTGTGTTT 1858  
Qy 1856 CCACCTTTGGAAACTCGGCAATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1915  
Db 1859 CCACCTTTGGAAACTCGGCAATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1917  
Qy 1916 CGATGAGTCGATGTTGGGTTTTTCCATTTTCAGGAAAGGATTTAGTGACAG-AATCAAGC 1974  
Db 1918 CGATGAGTCGATGTTGGGTTTTTCCATTTTCAGGAAAGGATTTAGTGACAGAAATCAAGC 1977  
Qy 1975 GAAAGAAATATTATTATGATCGGTGTTCCAGCATATAGGAAATTCAGGGGTTTAAAG 2034  
Db 1978 GAAAGAAATATTATTATGATCGGTGTTCCAGCATATAGGAAATTCAGGGGTTTAAAG 2037  
Qy 2035 TATAGAAATTTATTCGTAATTCAGTAGTATGTTGAAATTTGGAGGTTGCAATTTTGTCT 2094  
Db 2038 TATAGAAATTTATTCGTAATTCAGTAGTATGTTGAAATTTGGAGGTTGCAATTTTGTCT 2097  
Qy 2095 CACACATGTTATTTAAATTTTCTCATGTGCAATGGTGTGAGGATGACAGAAACAA 2154  
Db 2098 CACACATGTTATTTAAATTTTGTCTATGTGCA--TGTGTTGAGGATGAGC-AGAAACAA 2154  
Qy 2155 CAGGGCGGTTTCGTGGTTTTTGTCTGATGTGAAACTTTTTTAAAAATAACCAACAGAAA 2214  
Db 2155 CAGGGCGGTTTCGTGGTTTTTGTCTGATGTGCAATTTTTTAAAAATAACCAACAGAAA 2214  
Qy 2215 GGTGCAACTTTATGATGATAAGGTTTGGAAATGTCACCTCGGAATTAACCAAGAA 2274  
Db 2215 GGTGCAACTTTATGATGATAAGGTTTGGAAATGTCACCTCGGAATTAACCAAGAA 2274  
Qy 2275 GATAGAGTATTACTCCTTGTGTTCAATTTCACTACTAAACCAAAATCGATCTGTATGGT 2334  
Db 2275 GATAGAGTATTACTCCTTGTGTTCAATTTCACTACTAAACCAAAATCGATCTGTATGGT 2334  
Qy 2335 TATGTAACCTGTGACACAAATTTCAAAATTAACCTTGTCTGTTGGCAATGCTCTG 2394



Db 2335 TATGTAACTGTGTGACACAAATAATTTCAAATAAACTTGTCTGTGTGGCAATGCTCTG 2394  
Qy 2395 CTTTCAGCAATGCACTACTCTAAATTTTACTGGATCATCAACAAATCATTTATGGAAGGTTT 2454  
Db 2395 CTTTCAGCAATGCACTACTCTAA--TTTACTGGTCATCAACAAATCATTTATGGAAGGTTT 2452  
Qy 2455 ATTCTATTCTGTTCTTTTACGTTTCATGAAGTATAGCTAAATTTCAAACTCTCTGCACTACC 2514  
Db 2453 ATTCTATT--GTCTTTTACTTTCATGAAGTATAGCTAAATTTCAAACTCTCTGCACTACC 2510  
Qy 2515 TAAATTTAGTGGACCGAATCTAGTCATCGTCTCGCTTACGCTTACCTCCAAATTTGATGGCG 2574  
Db 2511 TAAATTTAGTGGACCGAATCTAGTCATCGTCTCTCTTACGCTATCTCCAAATTTGATGGCT 2570  
Qy 2575 TGTTCTAATTCATGATGCTGCTACTGAGA--ACTGGTGTGTTTAAACAGCAATGTAAGAATG 2633  
Db 2571 TGTTCTAATTCATGATGCTGCTACTGAGACACTGGTGTGTTTAAACAGCAATGTAAGAATG 2630  
Qy 2634 CCAG--TACACCCCTAAATGTTTATGAGGATAGGATAGTG-----AC 2673  
Db 2631 CCAGTTACACCCCTAAATGTTTATGAGGATAGGATAGTGTTGCAACCTTGATCCATAG 2690  
Qy 2674 AGTGTAACTAGAGCAAGTGTATGCAATTTGATGCTTTAGATAGTTCATAGCATCTGGCAG 2733  
Db 2691 AGGCAAAAGTGTATGCAAAATTTTATGATGCTTTTATGATAGTATTTAATAGGCATCTGGGCAGA 2750  
Qy 2734 AAGCTTGAATATTGATG--TCTAGAATGGTAAGACAATCTCTCGTGTCTTAAATAAAAAA 2792  
Db 2751 AAGCTTGAATATTGATGTTCTAGAATGGTAAGACAATCTCTCGTGTCTTAAATAAAAAA 2810  
Qy 2793 AAGGTAAAGACATCAATTTTGTACATCCCTTTATTTTACTAAATTTTGTGTCACCTGGCT 2852  
Db 2811 AAGGTAAAGACATCAATTTTGTACATCCCTTTATTTTACTAAATTTTGTGTCACCTGGCT 2870  
Qy 2853 GCGGAAGGGAACACAGTCAACCTGATTAAGATGAATTTTGTGTCACCTTTAGCCA 2912  
Db 2871 GCGGAAGGGAACACAGTCAACCTGATTAAGATGAATTTTGTGTCACCTTTAGCCA 2930  
Qy 2913 CTGCTGATATGTTGAGGGCTGCACTGGCTGCTAAACTCCACTTGGGATTAAGGGCTAAAG 2972  
Db 2931 CTGCTGATATGTTGAGGGCTGCACTGGCTGCTAAACTCCACTTGGGATTAAGGGCTAAAG 2990  
Qy 2973 AAGCTATGGACAGGTAGTTTTTAAAGAACATATAGCAACAGNAATTAACACGACAGA 3032  
Db 2991 AAGCTATGGACAGGTAGTTTTTAAAGAACATATAGCAACAGNAATTAACACGACAGA 3050  
Qy 3033 ATGGGTTTCTTGATCTTTTGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTGATG 3092  
Db 3051 ATGGGTTTCTTGATCTTTTGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTGATG 3110  
Qy 3093 ACTTGTTGTTGGGATTTATGATGAAGCCATGAAGAAAATCTTCATGCCAGAAAGGTTTAA 3152  
Db 3111 ACTTGTTGTTGGGATTTATGATGAAGCCATGAAGAAAATCTTCATGCCAGAAAGGTTTAA 3170  
Qy 3153 TCCTTGATGTTTCCCTPAGAACTGTTGTTCAAGCACAGAGGTGAGGTCCTTGGTCAATA 3212  
Db 3171 TCCTTGATGTTTCCCTPAGAACTGTTGTTCAAGCACAGAGGTGAGGTCCTTGGTCAATA 3230  
Qy 3213 TGCACCGCTATATAAAAGAGCTCCTTTTGTGTTTATGAGCTGCTATATAAATGGACAGT 3272  
Db 3231 TGCACCGCTATATAAAAGAGCTCCTTTTGTGTTTATGAGCTGCTATATAAATGGACAGT 3290  
Qy 3273 TTCTATCATTTGATCACTTTTCTTACTTAAATAATGGTGGAGCTTGATGAATGTTGGCCA 3332  
Db 3291 TTCTATCATTTGATCACTTTTCTTACTTAAATAATGGTGGAGCTTGATGAATGTTGGCCA 3350  
Qy 3333 AACAAAGTACTAAGATTGACAAGGTTCTTAAATTTTCAATTTGATGATGAATCTGGAAG 3392  
Db 3351 AACAAAGTACTAAGATTGACAAGGTTCTTAAATTTTCAATTTGATGATGAATCTGGAAG 3410  
Qy 3393 AACGAATACCGTCTGTTGATGATCCACCCATCAAGCTGGTAGATCTTATCATACAAATTTG 3452  
Db 3411 AACGAATACCGTCTGTTGATGATCCACCCATCAAGCTGGTAGATCTTATCATACAAATTTG 3470

Qy 3453 CTCTCTCTAAGACTCTGCACTTGATGATGAAGTATACCAAGATTAATGCTCTCGCTT 3512  
Db 3471 CTCTCTCTAAGACTCTGCACTTGATGATGAAGTATACCAAGATTAATGCTCTCGCTT 3530  
Qy 3513 GCATTTGTGAGTACATCAGATTTTACCATTTTTCATTTTCTAAATAGATTGGTACA 3572  
Db 3531 GCATTTGTGAGTACATCAGATTTT--ACCATTTTTCATTTTCTATAGATTGGTACA 3588  
Qy 3573 TATAATGGTTGATGCTTGCAGCATATGCGCCTTTTACCATAATTTCCCTGTCTCATCTCG 3632  
Db 3589 TATAATGGTTGATGCTTGCAGCATATGCGCCTTTTACCATAATTTCCCTGTCTCATCTCG 3648  
Qy 3633 ATCATCGTAGTACTGCTTCTTCTTTTAAAGCAAAACATGAGCCATCATTTCTTTC 3692  
Db 3649 ATCATCGTAGTACTGCTTCTTCTTTTAAAGCAAAACATGAGCCATCATTTCTTTC 3708  
Qy 3693 AAAAAATAAAAAAGGTTTCATTGAGTACCTTCTCTTCTGCCACTGTTTATTTAGTTGCTTCC 3752  
Db 3709 -AAAAATAAAAAAGGTTTCATTGAGTACCTTCTCTTCTGCCACTGTTTATTTAGTTGCTTCC 3767  
Qy 3753 CAATAGTTTAAATAGTAGTATGATATGATATATTTTCTTGTTTTGGTTGTTCTCG 3812  
Db 3768 CAATAGTTTAAATAGTAGTATGATATGATATATTTTCTTGTTTTGGTTGTTCTCG 3827  
Qy 3813 TACCAAGAGTAAATGCACCTTTTGTGACAAGAAATGATAGGCAGTGTTTGCATGACAC 3872  
Db 3828 TACCAAGAGTAAATGCACCTTTTGTGACAAGAAATGATAGGCAGTGTTTGCATGACAC 3887  
Qy 3873 ACTTTTGTCTCTTTTCTGACAAATTAATGCTGTTTAAAGTGGCCATAAATAGATACATCGA 3932  
Db 3888 ACTTTTGTCTCTTTTCTGACAAATTAATGCTGTTTAAAGTGGCCATAAATAGATACATCGA 3947  
Qy 3933 CATGTTTTTGTAGCAGGAGATTGATATTTGTTTCTATTGCTTCCAAATTTAAAGCATATT 3992  
Db 3948 CATG--TTTTTGTAGCAGGAGATTGATATTTGTTTCTATTGCTTCC--ATTAAAGCATATT 4005  
Qy 3993 CTTCTTTAGCAATGATTTTCATGTGGGACATATTTGCTGCTGCTATTAAAGTAAATTTGTTG 4052  
Db 4006 CTTCTTTAGCAATGATTTTCATGTGGGACATATTTGCTGCTGCTATTAAAGTAAATTTGTTG 4065  
Qy 4053 ATATCATATATATCTTTTAAATGTTAAATATTTGTCATCTGCTCTCTCTGATTTGCTTTG 4112  
Db 4066 ATATCATATATATCTTTTAAATGTTAAATATTTGTCATCTGCTCTCTCTGATTTGCTTTG 4125  
Qy 4113 TCTTTTCAAAAAGGTTTACTGAGAAACCTTTAAATTTCAAAAGGAAAGATGACAGCTGCACT 4172  
Db 4126 TCTTTTCAAAAAGGTTTACTGAGAAACCTTTAAATTTCAAAAGGAAAGATGACAGCTGCACT 4185  
Qy 4173 ATTGAAGTCAGGCTTGAAGCCTTCCACGTACAAACTAAGCCTGTATGTTTCTCTTTAGCA 4232  
Db 4186 ATTGAAGTCAGGCTTGAAGCCTTCCACGTACAAACTAAGCCTGTATGTTTCTCTTTAGCA 4245  
Qy 4233 ACTACGTTTTTAAATATTCAGATATTTCTTTTAAAGTATGTCGTACTTTCAGTTAAGGCGG 4292  
Db 4246 ACTACGTTTTTAAATATTCAGATATTTCTTTTAAAGTATGTCGTACTTTCAGTTAAGGCGG 4305  
Qy 4293 ATTCTTTCAGTTGCAATCAGTGTCTTCTGATATATCTTTTCAATTTGTTTTT 4341  
Db 4306 ATTCTTTCAGTTGCAATCAGTGTCTTCTGATATATCTTTTCAATTTGTTTTT 4354

## RESULT 2

US-09-802-927-2  
; Sequence 2, Application US/09802927  
; Patent No. US20010031236A1  
; GENERAL INFORMATION:  
; APPLICANT: UCHIMIYA, HIROFUMI  
; APPLICANT: ARAI, SATOSHI  
; APPLICANT: FUSHIMI, TAKAOMI  
; APPLICANT: TAGAWA, MICHITO  
; APPLICANT: FUKUZAWA, HIROMITSU  
; TITLE OF INVENTION: DNA FRAGMENT HAVING PROMOTER FUNCTION

FILE REFERENCE: 204323USOCIP  
CURRENT APPLICATION NUMBER: US/09/802,927  
CURRENT FILING DATE: 2001-03-12  
PRIOR APPLICATION NUMBER: PCT/JP98/04088  
PRIOR FILING DATE: 1988-09-10  
PRIOR APPLICATION NUMBER: PCT/JP99/04847  
PRIOR FILING DATE: 1999-09-08  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2  
LENGTH: 1443  
TYPE: DNA  
ORGANISM: Oryza sativa L.cv.Nipponbare  
US-09-802-927-2

Query Match 31.9%; Score 1384.6; DB 10; Length 1443;  
Best Local Similarity 99.2%; Pred. No. 1.9e-281;  
Matches 1434; Conservative 0; Mismatches 4; Indels 7; Gaps 4;

QY 1 CTGCAGGAAGATTAAATAGGTGGACACACCAACCCCTGTGGTGGTGACGCCCTGTGTT 60  
DB 1 CTGCAGGAAGATTAAATAGGTGGACACACCAACCCCTGTGGTGGTGACGCCCTGTGTT 60  
QY 61 AATCAACTGGGGTTCGTTGGACATGTTTTCAGGAAATTAAGCAAGAAAATTAAG 120  
DB 61 AATCAACTGGGGTTCGTTGGACATGTTTTCAGGAAATTAAGCAAGAAAATTAAG 120  
QY 121 AAGATGCTCAAGCTGATGATGAGAAACGTAATCCATGGAAGCGAATTCAGTCGTT 180  
DB 121 AAGATGCTCAAGCTGATGATGAGAAACGTAATCCATGGAAGCGAATTCAGTCGTT 180  
QY 181 TCTTGTAACCTACCTGTTAGATACATAGACAGTGCACAC- GTTTGATGGCTCCTATTG 239  
DB 181 TCTTGTAACCTACCTGTTAGATACATAGACAGTGCACACGTTTGATGGCTCCTATTG 240  
QY 240 GCTCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 299  
DB 241 GCTCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
QY 300 ACCGAAACCCCAAGATTAATCTATTCCACCTCAGGGTAATGTGCTCAACTATGCAAT 359  
DB 301 ACCGAAACCCCAAGATTAATCTATTCCACCTCAGGGTAATGTGCTCAACTATGCAAT 360  
QY 360 GAATACAAATTCGCAAAATATCATGTTATCTATCTGTCTCAAAATGAAATTTGAGTCA 419  
DB 361 GAATACAAATTCGCAAAATATCATGTTATCTATCTGTCTCAAAATGAAATTTGAGTCA 420  
QY 420 ACTGAGACTCAATACGATTTTCTTTTCAAAAGAAATTTATTAATTTTTCATGAA 479  
DB 421 ACTGAGACTCAATACGATTTTCTTTTCAAAAGAAATTTATTAATTTTTCATGAA 480  
QY 480 ACGCAATTCACCGTTCGAGAAATATGCTGTCTATAAATAGTAGTCTAGTCGAGAAACAA 539  
DB 481 ACGCAATTCACCGTTCGAGAAATATGCTGTCTATAAATAGTAGTCTAGTCGAGAAACAA 540  
QY 540 AATTAATATCACATAAAGAGAGTTGTTAATTAACCACTGTTTCTGCTACTCAACTC 599  
DB 541 AATTAATATCACATAAAGAGAGTTGTTAATTAACCACTGTTTCTGCTACTCAACTC 600  
QY 600 TAATTTGTAATTTCTTATTTTTCAGTCACAAATTTCCAAATTTCCAAATTAAGAAAATAAACG 659  
DB 601 TAATTTGTAATTTCTTATTTTTCAGTCACAAATTTCCAAATTTCCAAATTAAGAAAATAAACG 660  
QY 660 TAGACGCTAAGCCCAACCTCTAAGCTAAGTTTCAGAGGTTGAAGTACGACGAGAAAT 719  
DB 661 TAGACGCTAAGCCCAACCTCTAAGCTAAGTTTCAGAGGTTGAAGTACGACGAGAAAT 720  
QY 720 ATGATGTTTATTAATATGATTTTTTTTAAATAACCTTACATAAATTTCTTTAGCAAC 779  
DB 721 ATGATGTTTATTAATATGATTTTTTTTAAATAACCTTACATAAATTTCTTTAGCAAC 780  
QY 780 ATATCATTTAATGTTTGAACCAACCTGTCACATAGAAAATTAAGAACGATGATGTTGGAA 839  
DB 781 ATATCATTTAATGTTTGAACCAACCTGTCACATAGAAAATTAAGAACGATGATGTTGGAA 838  
QY 840 ACAAGAGAAAAACACAGCCCTTAAGGCTTCTTGATCCTCTAGTTGGAGGTTGATTTTCAAA 899  
DB 839 ACAAGAGAAAAACACAGCCCTTAAGGCTTCTTGATCCTCTAGTTGGAGGTTGATTTTCAAA 898  
QY 900 CGCATGATAACGAGAAAGCTCAATAGCAATATTAATCTAGATATTTAATTAATTAAC 959  
DB 899 CGCATGATAACGAGAAAGCTCAATAGCAATATTAATCTAGATATTTAATTAATTAAC 958  
QY 960 TTGAAAAAATATTTATTTGAATTTTTTAAACAATGATATGATATTTTAAAAAA 1019  
DB 959 TTGAAAAAATATTTATTTGAATTTTTTAAACAATGATATGATATTTTAAAAAA 1018  
QY 1020 CACACCAATTTAACCTTTTAAAGCAATCTTAATAGGAACGAGGAAGTTAAAGATTAC 1079  
DB 1019 CACACCAATTTAACCTTTTAAAGCAATCTTAATAGGAACGAGGAAGTTAAAGATTAC 1078  
QY 1080 CGAA---GTGCTTTGGATAATGAAAAATGGGTGGGATTAAGAATTTGTAATCAATCAGGG 1136  
DB 1079 CGAAGGTGTGTTGGATAATGAAAAATGGGTGGGATTAAGAATTTGTAATCAATCAGGG 1138  
QY 1137 TTAGGATTAATAATTTAAATGAAAGAGGAGAAATGAATGGTTAGAGTTTAAATGTGCTT 1196  
DB 1139 TTAGGATTAATAATTTAAATGAAAGAGGAGAAATGAATGGTTAGAGTTTAAATGTGCTT 1198  
QY 1197 TTTGGTGGTAGAAAATTTTCCCATACACTCCCGAGAGCGGTGCGTTCGTTGGTGG 1256  
DB 1199 TTTGGTGGTAGAAAATTTTCCCATACACTCCCGAGAGCGGTGCGTTCGTTGGTGG 1258  
QY 1257 CAGAAGCGTCTTTTTCGTTGGAAAAAACTGCTTTAAAAAGGAAACAGAAAGAGAGAGAG 1316  
DB 1259 CAGAAGCGTCTTTTTCGTTGGAAAAAACTGCTTTAAAAAGGAAACAGAAAGAGAGAGAG 1318  
QY 1317 GTTT-GTTGTCCCGTCTCACCAGAAACGAAACCAAAAGCCCAACCTTAACCTCTC 1375  
DB 1319 GTTTGTGTTGTCACCGTCTCACCAGAAACGAAACCAAAAGCCCAACCTTAACCTCTC 1378  
QY 1376 GATCCGACCGAGACTCTCTCAATTTTCAGCGCGCAGCGGAGAGACGCGACGCGAGTTG 1435  
DB 1379 GATCCGACCGAGACTCTCTCAATTTTCAGCGCGCAGCGGAGAGACGCGACGCGAGTTG 1438  
QY 1436 TCGAC 1440  
DB 1439 TCGAC 1443

RESULT 3

US-09-878-574-3892  
Sequence 3892, Application US/09878574  
Patent No. US20020110548A1  
GENERAL INFORMATION:  
APPLICANT: Byrum, Joseph R.  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Thompson, Michael D.  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
FILE OF INVENTION: Plance  
FILE REFERENCE: 38-21(15401) B  
CURRENT APPLICATION NUMBER: US/09/878,574  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 09/333,535  
PRIOR FILING DATE: 1999-06-14  
NUMBER OF SEQ ID NOS: 15775  
SEQ ID NO 3892  
LENGTH: 404  
TYPE: DNA  
ORGANISM: Glycine max  
OTHER INFORMATION: Clone ID: LIB3028-004-Q1-B1-E12  
US-09-878-574-3892

Query Match 2.8%; Score 122; DB 10; Length 404;  
Best Local Similarity 80.3%; Pred. No. 2e-16; Indels 0; Gaps 0;  
Matches 143; Conservative 0; Mismatches 35;

Qy 3310 GCAGCTTGATGAATGTTGGCCAAACAAGGTACTAAGATTGACAAGGTTCTAAATTTTGC 3369  
Db 144 GAAGCTTGATGAGATGCTGCTCAAAACAAGGAGTTAAAGTTGATAAGGTCTCAATTTTGC 203  
Qy 3370 AATTGATGATGAATCTGGAAGAAGATTACCGGTGCTGGATCCACCCATCAAGTGG 3429  
Db 204 AATTGATGATGAATCTTGGAGAGCAATTACTGCTGCTGATACACCCATCCAGTGG 263  
Qy 3430 TAGATCTTATCATACAAAATTTGCTCTCTAAGACTCTCTGACTTGTGATGATGAAGT 3487  
Db 264 CAGAATCTACCATACAAAATTTTCCCTCCAAAGGTTCTTGGTGTGATGATGTTACT 321

RESULT 4

US-09-770-445-271/c  
; Sequence 271, Application US/09770445  
; Patent No. US20020023281A1  
; GENERAL INFORMATION:  
; APPLICANT: Gorlach, Jorn  
; APPLICANT: An, Yong-Qiang  
; APPLICANT: Hamilton, Carol M.  
; APPLICANT: Price, Jennifer L.  
; APPLICANT: Raines, Tracy M.  
; APPLICANT: Yu, Yang  
; APPLICANT: Rameaka, Joshua G.  
; APPLICANT: Page, Amy  
; APPLICANT: Matthew, Abraham V.  
; APPLICANT: Ledford, Brooke L.  
; APPLICANT: Woessner, Jeffrey P.  
; APPLICANT: Haas, William David  
; APPLICANT: Garcia, Carlos A.  
; APPLICANT: Kriker, Maja  
; APPLICANT: Slader, Ted  
; APPLICANT: Davis, Keith R.  
; APPLICANT: Allen, Keith  
; APPLICANT: Hoffman, Neil  
; APPLICANT: Hurban, Patrick  
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
; TITLE OF INVENTION: thaliana  
; FILE REFERENCE: 2023US (PARA-012PRV)  
; CURRENT APPLICATION NUMBER: US/09/770,445  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: US 60/178,472  
; NUMBER OF SEQ ID NOS: 999  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 271  
; LENGTH: 981  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(981)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-770-445-271

Query Match 2.6%; Score 112.8; DB 10; Length 981;  
Best Local Similarity 76.7%; Pred. No. 2.5e-14;  
Matches 138; Conservative 0; Mismatches 42; Indels 0; Gaps 0;  
Qy 3308 GTGCAGCTTGATGAATGTTGGCCAAACAAGGTACTAAGATTGACAAGGTTCTAAATTTT 3367  
Db 561 GAGAAGCTCGATGAGATGCTTAGAGCGGAGGAACTGAATTGACAAAGTTCTCACTTT 502  
Qy 3368 GCAATTGATGATCAATPACTGGAAGAACGAATTAACCGGTGTTGGATCCACCCATCAAGT 3427  
Db 501 GCTATTGATCAGCAATCTTTGGAGGAAGAATAACCGGGGATGATCCACCCATCGAGT 442  
Qy 3428 GGTAGATCTTATCATACAAAATTTGCTCTCTAAGACTCTCTGACTTGTGATGATGAAGT 3487  
Db 441 GCGAGGAGTTACCACACTAAATTTGCTCTCTCCCAAAACCCCTGGAGTTGATGATATTACT 382

RESULT 5

US-09-294-093B-1791  
; Sequence 1791, Application US/09294093B  
; Patent No. US20010051335A1  
; GENERAL INFORMATION:  
; APPLICANT: Lalgudi, Raghunath, V.  
; APPLICANT: Ito, Laura, Y.  
; APPLICANT: Sherman, Bradley, K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL  
; FILE REFERENCE: PL-0009 US  
; CURRENT APPLICATION NUMBER: US/09/294,093B  
; CURRENT FILING DATE: 1999-04-16  
; PRIOR APPLICATION NUMBER: 60/082,567  
; PRIOR FILING DATE: April 21, 1998  
; NUMBER OF SEQ ID NOS: 6207  
; SOFTWARE: PERL Program  
; SEQ ID NO 1791  
; LENGTH: 295  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700344966H1  
US-09-294-093B-1791

Query Match 2.4%; Score 102.2; DB 10; Length 295;  
Best Local Similarity 84.1%; Pred. No. 2.6e-12;  
Matches 127; Conservative 0; Mismatches 23; Indels 1; Gaps 1;  
Qy 2838 TAGTCCACCTGCTGCGGAAAGGAAACAGTCACCGCTGATTAGGATGAATTTTGC 2896  
Db 144 TCGGCCACCTGCTGCGGAAAGGTACTCAGTCTCTTATTAGGATGAATTTGC 203  
Qy 2897 TTGTGCCATTAGCCACTGGTGATATGTTGAGGCTGCTGCTAAACTCCACTT 2956  
Db 204 CTGTGCCATTAGCCACTGGTGATATGTTGAGGCTGCTGCTAAACTCCACTT 2956  
Qy 2957 GGGATTAAAGCTAAAGAAAGCTATGGACAAGG 2987  
Db 264 GGTATCAAGCTAAAGAAAGCTATGGATAAGG 294

RESULT 6

US-09-770-445-484/c  
; Sequence 484, Application US/09770445  
; Patent No. US20020023281A1  
; GENERAL INFORMATION:  
; APPLICANT: Gorlach, Jorn  
; APPLICANT: An, Yong-Qiang  
; APPLICANT: Hamilton, Carol M.  
; APPLICANT: Price, Jennifer L.  
; APPLICANT: Raines, Tracy M.  
; APPLICANT: Yu, Yang  
; APPLICANT: Rameaka, Joshua G.  
; APPLICANT: Page, Amy  
; APPLICANT: Matthew, Abraham V.  
; APPLICANT: Ledford, Brooke L.  
; APPLICANT: Woessner, Jeffrey P.  
; APPLICANT: Haas, William David  
; APPLICANT: Garcia, Carlos A.  
; APPLICANT: Kriker, Maja  
; APPLICANT: Slader, Ted  
; APPLICANT: Davis, Keith R.  
; APPLICANT: Allen, Keith  
; APPLICANT: Hoffman, Neil  
; APPLICANT: Hurban, Patrick  
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
; TITLE OF INVENTION: thaliana  
; FILE REFERENCE: 2023US (PARA-012PRV)  
; CURRENT APPLICATION NUMBER: US/09/770,445  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: US 60/178,472

;; PRIOR FILING DATE: 2000-01-27  
;; NUMBER OF SEQ ID NOS: 999  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 484  
;; LENGTH: 894  
;; TYPE: DNA  
;; ORGANISM: Arabidopsis thaliana  
;; NAME/KEY: misc feature  
;; FEATURE:  
;; LOCATION: (1)...(894)  
;; OTHER INFORMATION: n = A,T,C or G

US-09-770-445-484

Query Match 2.3%; Score 101.8; DB 10; Length 894;  
Best Local Similarity 71.9%; Pred. No. 4.9e-12;  
Matches 133; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy 3303 AATGGTCAGCTGATGAATGTTGGCCAAACAAGGTACTAAGATTGACAAAGGTTCTAA 3362

Db 583 AAGCAGAGAAGCTTGTGATGCTTAATAGAAAGGGAGCTCAGATAGATAAGGTGCTTA 524

Qy 3363 ATTTGGCAATTGATGATCAATACACTGGAGACGAATTACCGTCTGTTGGATCCACCAT 3422

Db 523 ATTTGGCAGTGATGATTCGGTCTCGAAGAAAGAAATTACTGGAAGGTGGATTCACCTT 464

Qy 3423 CAAGTGGTAGATCTTATCATACAAAATTGCTCTCTTAAGACTCTCTGACTTGATGATG 3482

Db 463 CAAGTGGAGAAGCTATCATACTAAATTTCGCACCTCTTAAGTTCCAGGAGTCGATGATC 404

Qy 3483 TAAAT 3487

Db 403 TGAAT 399

#### RESULT 7

US-09-923-876-3999  
; Sequence 3999, Application US/09923876  
; Patent No. US20020013958A1

;; GENERAL INFORMATION:  
;; APPLICANT: Lalgudi, Raghunath V.  
;; APPLICANT: Kamigaki, Laura Y. (Ito)  
;; APPLICANT: Sherman, Bradley K.  
;; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING  
;; FILE REFERENCE: PL-0012-1 CON  
;; CURRENT APPLICATION NUMBER: US/09/923,876  
;; CURRENT FILING DATE: 2001-08-06  
;; PRIOR APPLICATION NUMBER: 09/298,329  
;; PRIOR FILING DATE: 1999-04-21  
;; PRIOR APPLICATION NUMBER: 60/085,331  
;; PRIOR FILING DATE: 1998-05-05  
;; NUMBER OF SEQ ID NOS: 6332  
;; SOFTWARE: PERL Program  
;; SEQ ID NO 3999  
;; LENGTH: 292  
;; TYPE: DNA  
;; ORGANISM: Zea mays  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; OTHER INFORMATION: Incyte ID No. US20020013958A1 700454231H1

;; LOCATION: 34, 59, 83, 88, 92, 113, 122, 142, 145, 158, 172-173, 217, 236-237, 255,  
;; LOCATION: 260, 277-278, 284  
;; OTHER INFORMATION: a, t, c, g, or other

US-09-923-876-3999

Query Match 2.1%; Score 92.8; DB 10; Length 292;  
Best Local Similarity 73.2%; Pred. No. 2.5e-10;  
Matches 123; Conservative 0; Mismatches 44; Indels 1; Gaps 1;

Qy 2838 TAGGTCCACCTGGCTGGCGAAGGGAACACAGTCAACGCTGATTAAGGATGAATTTGCT 2897

Db 109 TCGNCCACCTGGTCTGTGAAAGGTACTTCAGTNCCTTATTAAGGAATATTGCC 168

Qy 2898 TGTCACATTTAGCCACTGGTGATATGTTGAGGGTGCAGTGGCTGCTAAAACTCCACTTG 2957  
Db 169 TGTNNCAITTTAGCCACTGGTGATATGTTGAGGGTGCAGTGGCTGCTGACGCCANGAC-AACTCTAG 227  
Qy 2958 GGATTAAGGCTAAAGAAGCTATGGACAAGGTAGTTTTTTAAGAAACATA 3005  
Db 228 GTATCAAAANNTAAAGAAGCTATGGATANGGANAGCTTTGTTTCAGATA 275

#### RESULT 8

US-09-923-876-6281  
; Sequence 6281, Application US/09923876  
; Patent No. US20020013958A1

;; GENERAL INFORMATION:  
;; APPLICANT: Lalgudi, Raghunath V.  
;; APPLICANT: Kamigaki, Laura Y. (Ito)  
;; APPLICANT: Sherman, Bradley K.  
;; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING  
;; FILE REFERENCE: PL-0012-1 CON  
;; CURRENT APPLICATION NUMBER: US/09/923,876  
;; CURRENT FILING DATE: 2001-08-06  
;; PRIOR APPLICATION NUMBER: 09/298,329  
;; PRIOR FILING DATE: 1999-04-21  
;; PRIOR APPLICATION NUMBER: 60/085,331  
;; PRIOR FILING DATE: 1998-05-05  
;; NUMBER OF SEQ ID NOS: 6332  
;; SOFTWARE: PERL Program  
;; SEQ ID NO 6281  
;; LENGTH: 267  
;; TYPE: DNA  
;; ORGANISM: Zea mays  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; OTHER INFORMATION: Incyte ID No. US20020013958A1 700458807H1

;; NAME/KEY: unsure  
;; LOCATION: 47, 77, 98, 174  
;; OTHER INFORMATION: a, t, c, g, or other

US-09-923-876-6281

Query Match 2.1%; Score 92; DB 10; Length 267;  
Best Local Similarity 81.9%; Pred. No. 3.5e-10;  
Matches 104; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 1466 GCTGCGGAGATGGCGGCGAACTGGAGGACGTGGCGTGCATGGAGCTGATGACGAGCTG 1525

Db 42 GCGGTNGAGATGGCGGCGAACTGGAGGACGTGCANTCACTGGATCTGATGACGAGCTG 101

Qy 1526 CTCGCCCGCATGAAGTGCAGCTCCAGCCGACAGCGCGTCACTCGTGGTAACGCC 1585

Db 102 CTCGCCCGCATGAAGTGCAGCTCCAGCCGACAGCGCGTCACTCGTGGTAACGCC 161

Qy 1586 CGCTCT 1592

Db 162 TGGCTCT 168

#### RESULT 9

US-09-887-576-828/c  
; Sequence 828, Application US/09887576  
; Patent No. US2002014047A1

;; GENERAL INFORMATION:  
;; APPLICANT: Budworth, P.  
;; APPLICANT: Brown, D.  
;; APPLICANT: Chang, H.  
;; APPLICANT: Zhu, T.  
;; APPLICANT: Han, B.  
;; APPLICANT: Wang, X.  
;; APPLICANT: Cooper, Bret  
;; TITLE OF INVENTION: Promoters for regulation of plant expression  
;; FILE REFERENCE: 1360.001US1  
;; CURRENT APPLICATION NUMBER: US/09/887,576  
;; CURRENT FILING DATE: 2001-06-25  
;; PRIOR APPLICATION NUMBER: US 60/213,848

```
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 828
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
; US-09-887-576-828

Query Match      1.4%; Score 62.4; DB 10; Length 2000;
Best Local Similarity 61.9%; Pred. No. 0.0013;
Matches 99; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY  903 ATGATACAGCAAGCTCATTAGCACATTATTACTAGATATTTATTAATAAATCTG 962
Db  1227 ACGTAAACAGGATGAGCCATTACGTATGATTAATTAAGTATTAATCTTTTAAACTTT 1168
QY  963 AAAAAATATTATTGTAATTTTAAACCAATGTATGCATAAAATTTTAAAAAACAC 1022
Db  1167 AAAATAGATTAATATGATTTTAAAGCAACTTCTATAGAAATTTTGGAAAAAC 1108
QY  1023 ACCAATTTAACCTTTAAAAAGCATCTCTAATAGGAACGA 1062
Db  1107 ACACCGTTTAGTACTCGGAACGTCGCGGTGGAATGA 1068

RESULT 10
US-09-815-242-6911
; Sequence 6911, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6911
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(645)
; US-09-815-242-6911

Query Match      1.4%; Score 57; DB 10; Length 639;
Best Local Similarity 59.6%; Pred. No. 0.011;
Matches 96; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY  2834 ATTTTAGTCCACCTGGCTGGGAAAGGAAACACAGTCACCGCTGATTAAGATGAATTT 2893
Db  13 ATTTAGGGCTTACCTGGTGCAGGTAAGGNACTCAAGCAAAAAATCGTAGAACATTC 72
QY  2894 TGCCTTGTCACATTTAGCCACTGGTGATATTTTGGGGCTGCAGTGGCTGCTAAAACTCCA 2953
Db  73 CATGTTGCACATATCTCAACAGGTGATATGTTCCGGCTGCATGGCAAACTCAAACTGAA 132
QY  2954 CTTGGGATTAAGGCTAAAGAAAGCTATGGACAAGGTAGTGT 2994

Query Match      1.4%; Score 61.6; DB 10; Length 645;
Best Local Similarity 59.9%; Pred. No. 0.0012;
Matches 106; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY  2835 TTTTAGTCCACCTGGCTGGGAAAGGAAACACAGTCACCGCTGATTAAGGATGAATTTT 2894
Db  14 TTTTAGTGCACCGGTGCAGGTAAGGCACTCAAGCAATTTATTATGACAAATTTG 73
QY  2895 GCTTGTGCCATTTAGCCACTCGTGATATGTTGAGGGCTGCAGTGGCTGCTAAAACTCCAC 2954
Db  74 GTATCCCGCAAAATTTCAACTCGTGATATGTTCCGTGCTGCAATCAAAGCGGGGACTGAAC 133
QY  2955 TTGGGATTAAGCTAAAGAAAGCTATGGACAAGGTAGTGTAAAGAAACATATAGCAACAG 3014
Db  134 TTGGCAAAACAGCTAAAGCATTAATGGATGAAGGTAATTAAGTCCAGATGAATTTAACCG 193

RESULT 11
US-09-815-242-9121
; Sequence 9121, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9121
; LENGTH: 639
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(639)
; US-09-815-242-9121
```



```

; APPLICATION NUMBER: US/08/910,386A
; FILING DATE: 13-AUG-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058950US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13341 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Oryza longistaminata
; STRAIN: IRBB21
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 11
; MAP POSITION: 11q, RG103
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2367..4205
; OTHER INFORMATION: /product= "receptor kinase-like protein"
; OTHER INFORMATION: /note= "Xa21 gene family member D"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 4201..9071
; OTHER INFORMATION: /note= "retrofit, a copia-like,
; OTHER INFORMATION: transposon-like element"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4484..8821
; OTHER INFORMATION: /product= "retrofit"
; OTHER INFORMATION: /gene= "gag/pol"
; FEATURE:
; NAME/KEY: intron
; LOCATION: 9915..11712
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 10020..10975
; OTHER INFORMATION: /note= "Krispie, transposon-like
; OTHER INFORMATION: element"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 12626..12750
; OTHER INFORMATION: /note= "Pop-012, transposon-like
; OTHER INFORMATION: element"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 13040..13248
; OTHER INFORMATION: /note= "Ds-rice2, transposon-like
; OTHER INFORMATION: element"
;
; US-08-910-386A-1
;
; Query Match 1.3%; Score 55.8; DB 8; Length 13341;
; Best Local Similarity 57.0%; Pred. No. 0.065;
; Matches 102; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
;
; QY 891 ATTTTCAACGATGATTAACGAGAAAGCTCATTAGACATTATTACTTAGATATTATA 950
; Db 1842 AATTGACAAGATCGTGAACGAAATAGCTGTAGTGATATAATTAAATTAAGTAGTAATT 1783
; QY 951 ATTATAAATCTGAAAAAATATTATTATTGAATTTTTTAAAAACAATGTATGCATAAAATTATT 1010
; Db 1782 ATTTTAAATCTTAAAAAATAGATTAATATATATTTTAAAGAAACTTCTCTATAGAAAGTTT 1723
; QY 1011 TTTTAAAAACACCAATTTTAAACCCCTTTAAAAAGCATCTCTAATAGGAAACGAGGAAGTT 1069
; Db 1722 TTTTCAAAAAACACACCACTTTAAACAGTTCTGAAAAACATACCGTGAAAAACGATGCTTTT 1664
;
; RESULT 15
; US-08-910-386A-6/c
; Sequence 6, Application US/08910386A
; Patent No. US20020092041A1
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Hulbert, Scot
; APPLICANT: Richter, Todd
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,386A
; FILING DATE: 13-AUG-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058950US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19639 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Oryza longistaminata
; STRAIN: IRBB21
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 11
; MAP POSITION: 11q, RG103
; FEATURE:
; NAME/KEY: -
; LOCATION: 5213..18201
; OTHER INFORMATION: /note= "Xa21 gene"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(5213..7889, 8732..9132)
; OTHER INFORMATION: /product= "receptor kinase-like protein"
; OTHER INFORMATION: /note= "Xa21 disease resistance gene"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 9645..9769
; OTHER INFORMATION: /note= "Pop-Oil1, transposon-like
; OTHER INFORMATION: element"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 13040..13248
; OTHER INFORMATION: /note= "Ds-rice1, transposon-like
; OTHER INFORMATION: element"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(15118..17720, 17827..18204)
```

; OTHER INFORMATION: /product= "receptor kinase-like protein"  
; OTHER INFORMATION: /note= "x21 gene family member C; 2 bp  
; OTHER INFORMATION: deletion causing frame-shift mutation of  
; OTHER INFORMATION: ORF compared to family member A1"  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 16183..16184  
; OTHER INFORMATION: /note= "location of 2 bp deletion  
; OTHER INFORMATION: compared to family member A1"  
; US-08-910-386A-6

Query Match 1.3%; Score 55.8; DB 8; Length 19639;  
Best Local Similarity 57.0%; Pred. No. 0.076;  
Matches 102; Conservative 0; Mismatches 77; Indels 0; Gaps 0;  
Qy 891 ATTTTCAACGCATCATAAACGAGAAAGCTCATTAGCACATTATTACTTAGATATTATA 950  
Db 4688 AATTGACAGATACGTGACGAGAAATAGCTGTTAGTGATTAATTAAGTAGTAATT 4629  
Qy 951 ATTATAAACTTGAAAAAATATTTTGAATTTTTTAAACAATCTATGCATAAAATTATT 1010  
Db 4628 ATTTTAAACTTAAAAAATAGATTAATATATTTTAAAGAACTTTTCTATAGAAAGTTT 4569  
Qy 1011 TTTTAAAAACACCAATTTAACCCCTTTAAAGAGCATCCTAATAGAAACGAGGAGTT 1069  
Db 4568 TTTCAAAAAACACACCACTTTAACAGTTCTGAAAAACATACGGGTGAAAAACGATGCTTTT 4510

Search completed: March 15, 2003, 20:31:01  
Job time : 469.605 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2003, 09:16:40 ; Search time 3558.28 Seconds  
(without alignments)  
11777.621 Million cell updates/sec

Title: US-09-802-937-2  
Perfect score: 1440  
Sequence: 1 ctgcagagattaattagg.....cgcgacgagtgctgcgac 1440

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pin:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
2	1373.6	95.4	98782	2	CNS08085	AL731756 Oryza sat
c 3	115.6	8.0	137348	8	AP003297	AP003297 Oryza sat
c 4	109.2	7.6	124815	2	AP005594	AP005594 Oryza sat
c 5	109.2	7.6	138025	8	AP003261	AP003261 Oryza sat
c 6	109.2	7.6	169663	2	CNS07YFX	AL713952 Oryza sat
c 7	108.8	7.6	126659	8	OSJN00015	AL606455 Oryza sat
c 8	108.8	7.6	155582	2	OSJN00094	AL606649 Oryza sat
c 9	108.4	7.5	100419	2	OSIG00047	AL732355 Oryza sat
c 10	107.6	7.5	100635	8	AP003446	AP003446 Oryza sat
c 11	107.6	7.5	144074	8	AP003302	AP003302 Oryza sat
c 12	106.4	7.4	156054	8	AB023482	AB023482 Oryza sat
c 13	104.8	7.3	129838	8	AC079038	AC079038 Oryza sat
c 14	103.6	7.2	170025	2	OSJN00031	AC079038 Oryza sat
c 15	102	7.1	131983	8	AC023240	AC023240 Oryza sat
c 16	101	7.0	143959	2	AC090055	AC090055 Oryza sat
c 17	100	6.9	6735	8	RIC3H3M	L28995 Oryza sativ
c 18	100	6.9	129845	2	AP004165	AP004165 Oryza sat
c 19	99.4	6.9	121615	2	AP003981	AP003981 Oryza sat
c 20	99.4	6.9	146670	8	AP003250	AP003250 Oryza sat
c 21	99	6.9	142596	8	AP003314	AP003314 Oryza sat
c 22	99	6.9	166753	8	AP002866	AP002866 Oryza sat
c 23	98.4	6.8	159893	8	OSJN00095	AL606652 Oryza sat
c 24	98.4	6.8	189834	8	OSJN00171	AL662974 Oryza sat
c 25	96.8	6.7	108805	2	CNS080CAF	AL731875 Oryza sat
c 26	96.8	6.7	150238	2	CNS080CA9	AL831796 Oryza sat
c 27	96.6	6.7	151100	8	AP003453	AP003453 Oryza sat
c 28	96.2	6.7	136553	2	AC093953	AC093953 Oryza sat
c 29	95.8	6.7	132407	8	OSJN00117	AL607002 Oryza sat
c 30	95.2	6.6	139566	2	AP004161	AP004161 Oryza sat
c 31	95.2	6.6	147954	2	AP003983	AP003983 Oryza sat
c 32	95.2	6.6	165394	8	AC025296	AC025296 Oryza sat
c 33	94.8	6.6	160174	2	AP005103	AP005103 Oryza sat
c 34	94.8	6.6	162010	2	AP004121	AP004121 Oryza sat
c 35	94.6	6.6	102119	2	AP003919	AP003919 Oryza sat
c 36	93.2	6.5	122559	2	AP004301	AP004301 Oryza sat
c 37	92.8	6.4	168698	2	OSJN00023	AL606588 Oryza sat
c 38	92.4	6.4	148985	8	AP002094	AP002094 Oryza sat
c 39	92.2	6.4	135874	8	AC113337	AC113337 Genomic s
c 40	92	6.4	142776	2	AP005109	AP005109 Oryza sat
c 41	92	6.4	167110	2	AC120505	AC120505 Oryza sat
c 42	91.4	6.3	1154	8	RICADKA	D10334 Oryza sativ
c 43	91.4	6.3	133595	2	AC129227	AC129227 Oryza sat
c 44	91.4	6.3	164679	8	AC078944	AC078944 Genomic S
c 45	91	6.3	137462	8	AP002538	AP002538 Oryza sat

ALIGNMENTS

RESULT 1  
AB041773  
LOCUS AB041773 4341 bp DNA linear PLN 20-APR-2000  
DEFINITION Oryza sativa Adk-a gene for adenylyate kinase, partial cds.  
ACCESSION AB041773  
VERSION AB041773.1 GI:7630192  
KEYWORDS adenylyate kinase.  
SOURCE Oryza sativa DNA.  
ORGANISM Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
REFERENCE 1 (bases 1 to 4341)  
AUTHORS Fukuzawa,H., Uchimiya,H., Tagawa,M. and Arai,S.  
TITLE Oryza sativa Adk-a gene, promoter region and partial ORF

JOURNAL Published Only in DataBase (2000)  
REFERENCE 2 (bases 1 to 4341)  
AUTHORS Fukuzawa,H., Uchimiya,H., Tagawa,M. and Arai,S.  
TITLE Direct Submission  
JOURNAL Submitted (12-APR-2000) Hiromitsu Fukuzawa, Nissan Chemical Industries, LTD, Shiraoka Research Station of Biological Science, 1470 Shiraoka, Minamisaitama, Saitama 349-0294, Japan  
(E-mail: fukuzawa@nissanchem.co.jp, Tel.: +81-480-92-2513)

FEATURES  
Location/Qualifiers  
1..4341  
/organism="Oryza sativa"  
/db\_xref="taxon:4530"  
Join(1475..1577,2841..2986,3074..3193,3314..3481,  
4126..4215)  
/gene="Adk-a"  
Join(1475..1577,2841..2986,3074..3193,3314..3481,  
4126..4215)  
/gene="Adk-a"  
/codon\_start=1  
/product="adenylate kinase"  
/protein\_id="BAA94761.1"  
/db\_xref="GI:7630193"  
/translation="MANLEDVPSMELTELLRRMKSSKEDKRVILVGPCCGKGTQ  
SPLIDFELCLGATKMLRAAIAAKTPLIGIKAKEMDKGELVDLVLVGIIDEAMKK  
TSCQKGFILDGPRVTVOQKLDKMLAKGKIDKLVNFAIDDLLEIRITGRWIHPS  
SGRSYHTFAPPKTPGLDDVTGEPLIQKDDTAALVLSRLEAFHVQTRP"

BASE COUNT 1284 a 758 c 880 g 1419 t  
ORIGIN

Query Match 100.0%; Score 1440; DB 8; Length 4341;  
Best Local Similarity 100.0%; Pred. No. 2.1e-245;  
Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAGGAGATTAAATAGGTGGACACACCAACCCCTGTGTGGTGACGCGCTGTGT 60  
DB 1 CTGCAGGAGATTAAATAGGTGGACACACCAACCCCTGTGTGGTGACGCGCTGTGT 60  
QY 61 AATCAACTGGGGTGTTCGTGGACATGTTTTTGCAGGAAATTAAGCAAGAAATTAAG 120  
DB 61 AATCAACTGGGGTGTTCGTGGACATGTTTTTGCAGGAAATTAAGCAAGAAATTAAG 120  
QY 121 AAGATGCTCAGCTGACATGAGAAACCTTAATCCATGGAACGGAATTCAGTCGTTTC 180  
DB 121 AAGATGCTCAGCTGACATGAGAAACCTTAATCCATGGAACGGAATTCAGTCGTTTC 180  
QY 181 TCTGTGTAATACCATGTTTGAATACATAGACAGTGCCCAACGTTTGATGGCTCTATTGG 240  
DB 181 TCTGTGTAATACCATGTTTGAATACATAGACAGTGCCCAACGTTTGATGGCTCTATTGG 240  
QY 241 CTGCTGTGATGACTGTGTGCAAAAGCATCAATTCCTCTTGGAGTATCTTTATTA 300  
DB 241 CTGCTGTGATGACTGTGTGCAAAAGCATCAATTCCTCTTGGAGTATCTTTATTA 300  
QY 301 CGAAATCCCAAGAAATTAATCTTATTCACCTCAGGGTAATGTGCTGAACATGCAATG 360  
DB 301 CGAAATCCCAAGAAATTAATCTTATTCACCTCAGGGTAATGTGCTGAACATGCAATG 360  
QY 361 AATACAAATTCGCAAAATATCATGTTATCTATCTTGTCAAATTAATTTGAGTCCAA 420  
DB 361 AATACAAATTCGCAAAATATCATGTTATCTATCTTGTCAAATTAATTTGAGTCCAA 420  
QY 421 CTGAGACTGCAATACGATTTTCTTTCAAAAGAAATTAATTAATTTTTCATGAA 480  
DB 421 CTGAGACTGCAATACGATTTTCTTTCAAAAGAAATTAATTAATTTTTCATGAA 480  
QY 481 CGCAATTCACCGTTCGAGAAATATGCTGTCAATAAGTAGTGTAGTGCAAGAACAA 540  
DB 481 CGCAATTCACCGTTCGAGAAATATGCTGTCAATAAGTAGTGTAGTGCAAGAACAA 540  
QY 541 ATTAATATCACATAAAAAAGAGAGTTTAAATTAACAAACCATGTTTCGTAACACTCT 600  
DB 541 ATTAATATCACATAAAAAAGAGAGTTTAAATTAACAAACCATGTTTCGTAACACTCT 600

QY 601 AATTTGTAATTCCTTATTTTCAGTCACAAAATTCCTCAATTTCCAAATTAAGAAAAATAAACC 660  
DB 601 AATTTGTAATTCCTTATTTTCAGTCACAAAATTCCTCAATTTCCAAATTAAGAAAAATAAACC 660  
QY 661 AGACGGCTAAGCCCAACCCATCTAAGGCTAAGTTTCAGAGGTGAAGTACGACGAGAAAAATA 720  
DB 661 AGACGGCTAAGCCCAACCCATCTAAGGCTAAGTTTCAGAGGTGAAGTACGACGAGAAAAATA 720  
QY 721 TGATGCTTTTAAATATATGATTTTAAATTAACCTTTCACATATAATTTCTTTAGGAACA 780  
DB 721 TGATGCTTTTAAATATATGATTTTAAATTAACCTTTCACATATAATTTCTTTAGGAACA 780  
QY 781 TATCATTTAATGTTTGAAGAAACGTCACATAAGAAAACTAAGAACGATGAGTTGGGAAA 840  
DB 781 TATCATTTAATGTTTGAAGAAACGTCACATAAGAAAACTAAGAACGATGAGTTGGGAAA 840  
QY 841 CAAGAGAAAAACACAGCCTTAAGGCTTCTGATCCTCTAGTTGGAGGTTGATTTTCAAAAC 900  
DB 841 CAAGAGAAAAACACAGCCTTAAGGCTTCTGATCCTCTAGTTGGAGGTTGATTTTCAAAAC 900  
QY 901 GCATGATTAACGAGAAAGCTCATTAGCACATTTATTACTTTAGATATTTATAATTTATAA 960  
DB 901 GCATGATTAACGAGAAAGCTCATTAGCACATTTATTACTTTAGATATTTATAATTTATAA 960  
QY 961 TGAAGAAAAATATTTATTTGAAATTTTAAACCAATGTATGCATAAATTTTAAAAAAC 1020  
DB 961 TGAAGAAAAATATTTATTTGAAATTTTAAACCAATGTATGCATAAATTTTAAAAAAC 1020  
QY 1021 ACACCAATTTAACCTTTAAAAAGCATCTTAATAGGAACGAGGAAGTTAAAGATTCACC 1080  
DB 1021 ACACCAATTTAACCTTTAAAAAGCATCTTAATAGGAACGAGGAAGTTAAAGATTCACC 1080  
QY 1081 GAAGTGTTCGATATAATGAAATGCGGTGGGATTAGAATTTGTAATGTAATCAGCGGTAG 1140  
DB 1081 GAAGTGTTCGATATAATGAAATGCGGTGGGATTAGAATTTGTAATGTAATCAGCGGTAG 1140  
QY 1141 GATTAAATATTTAAATGAAAGAGGAGAAATGAATGGTTAGAGTTTAAATGTCTCTTTTG 1200  
DB 1141 GATTAAATATTTAAATGAAAGAGGAGAAATGAATGGTTAGAGTTTAAATGTCTCTTTTG 1200  
QY 1201 GTGGGTAGAAAATTTTGGCCATACACTCCCGAGAGCGGTGCGTTCGCTGGCGGAG 1260  
DB 1201 GTGGGTAGAAAATTTTGGCCATACACTCCCGAGAGCGGTGCGTTCGCTGGCGGAG 1260  
QY 1261 AGCGTCTTTTTCGTTGGAAAAAACTGCTTAAAAAGGAAACAGAAAAGAGCCAGCTTT 1320  
DB 1261 AGCGTCTTTTTCGTTGGAAAAAACTGCTTAAAAAGGAAACAGAAAAGAGCCAGCTTT 1320  
QY 1321 GTTGTACCGTCTCACCAGAAACGAAACAAAAAGCCCCACCCTTAAACCTCTCGATCC 1380  
DB 1321 GTTGTACCGTCTCACCAGAAACGAAACAAAAAGCCCCACCCTTAAACCTCTCGATCC 1380  
QY 1381 GACCGAGACTCTCTCCATTTTCAGCGGCGCACGCGGAGACGCGAGGTTCTGTCGAC 1440  
DB 1381 GACCGAGACTCTCTCCATTTTCAGCGGCGCACGCGGAGACGCGAGGTTCTGTCGAC 1440

RESULT 2  
CNS08C85  
LOCUS  
DEFINITION  
Oryza sativa chromosome 12 clone OJ1220\_D01, \*\*\* SEQUENCING IN  
PROGRESS \*\*\*, in ordered pieces.  
ACCESSION  
AL731756  
VERSION  
AL731756.3  
KEYWORDS  
HTG; HTGS PHASE2; HTGS\_ACTIVEFIN.  
SOURCE  
Oryza sativa.  
ORGANISM  
Oryza sativa.  
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
REFERENCE  
1 (bases 1 to 98782)  
AUTHORS  
Choisne,N., Orjeda,G., Cattolico,L., Demange,N., Wincker,P.,  
Segurens,B., Pelletier,E., Scarpelli,C., Salanoubat,M.,

Weisenbach, J. and Quetier, F.  
Oryza sativa chromosome 12 sequencing  
Unpublished  
2 (bases 1 to 98782)  
Genoscope.  
Direct Submission  
Submitted (09-JUL-2002) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)  
- Web : www.genoscope.cns.fr)  
On Jul 11, 2002 this sequence version replaced gi:20372820.  
IMPORTANT: This sequence is unfinished and does not necessarily  
represent the correct sequence.  
Work on the sequence is in progress and the release of this data is  
based on the understanding that the sequence may change as work  
continue. The sequence may be contaminated with foreign sequence  
from E.coli, yeast, vector, phage, etc.  
The nucleotide sequence of this BAC clone was generated by  
combining Monsanto and Genoscope sequencing data.  
\* NOTE: This is a 'working draft' sequence.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* by the accession number will be preserved.  
Location/Qualifiers  
1..98782  
/organism="Oryza sativa"  
/cultivar="Nipponbare"  
/sub\_species="japonica"  
/db\_xref="taxon:4530"  
/chromosome="12"  
/clone="OJ1220\_D01"  
/collection="Monsanto"

BASE COUNT 30065 a 19192 c 20278 g 29238 t 9 others

ORIGIN

Query Match 95.4%; Score 1373.6; DB 2; Length 98782;  
Best Local Similarity 99.2%; Pred. No. 6.2e-234;  
Matches 1434; Conservative 0; Mismatches 4; Indels 8; Gaps 5;

Qy 1 CTCGAGAGAGATTAAATTAGTGGACACACCAAAACCTGTGGTGGTGGAGCCCTGTGGTT 60  
Db 63789 CTCGAGAGAGATTAAATTAGTGGACACACCAAAACCTGTGGTGGTGGTGGTT 63848

Qy 61 AATCAACTCGGGTGTGGTTGGACATGGTTTTCAGAGAAAATTAAGCAAGAAAATTAG 120  
Db 63849 AATCAACTCGGGTGTGGTTGGACATGGTTTTCAGAGAAAATTAAGCAAGAAAATTAG 63908

Qy 121 AAGAAATGCTCAAGCTGACATGAGAAAACGTAAATCCAATGGAGCGAATTTCAAGTCGTT 180  
Db 63909 AAGAAATGCTCAAGCTGACATGAGAAAACGTAAATCCAATGGAGCGAATTTCAAGTCGTT 63968

Qy 181 TCTTGACTACTACCATGTTTGAATAATCAATGAAGCAAGTGGCAAC - GTTGTGATGGTCCATTG 239  
Db 63969 TCTTGACTACTACCATGTTTGAATAATCAATGAAGCAAGTGGCAACGGTTTGAATGGTCCATTG 64028

Qy 240 GCTCGTGTGATCTGACTGTGTGTCAAAAGCATCAAAATTCCTTCTGGAGTATCTTTATT 299  
Db 64029 GCTCGTGTGATCTGACTGTGTGTCAAAAGCATCAAAATTCCTTCTGGAGTATCTTTATT 64088

Qy 300 ACCGAAAAACCCCAAGATTATTTCTATTCCACCTCAGGGTAATTTGTGCTGAACATATGCAAT 359  
Db 64089 ACCGAAAAACCCCAAGATTATTTCTATTCCACCTCAGGGTAATTTGTGCTGAACATATGCAAT 64148

Qy 360 GAATACAAATTCGAAAATATCATGGTTATCTATCTTGTCTCAAAATTTGAAGTCA 419  
Db 64149 GAATACAAATTCGAAAATATCATGGTTATCTATCTTGTCTCAAAATTTGAAGTCA 64208

Qy 420 ACTGAGACTGCAATACGATTTTCTTTTCAAAAAGAAATTAATTAATTTTTTTTCATGAA 479  
Db 64209 ACTGAGACTGCAATACGATTTTCTTTTCAAAAAGAAATTAATTAATTTTTTTTCATGAA 64268

Qy 480 ACCGAATTCACCGTTTCGAGAAATATGCTGTCAATAATAAGTAGTCTAGTCGAGAAACAA 539  
Db 64269 ACCGAATTCACCGTTTCGAGAAATATGCTGTCAATAATAAGTAGTCTAGTCGAGAAACAA 64328



```

/db_xref="GI:20804752"
/translation="MSGSPFAAIAAAEIALASWANKLERDRVATRRWEGPAIAKPARTI
WGPPIRRSTAVATVTRTPDSPRRHGRSGRATRRMLSHLVLVSPANACRIAA
VNHARRIGNQKTRTTPRCCHAIRVWFMRLRAAAAAAASVLRQFTWMPOLCNQORAY
EQIQCYYVW"
complement(join(31579..31693,31909..32131,32235..32469,
32585..32716,32809..33174))
/gene="P0698A10.8"
CDS
complement(join(31579..31693,31909..32131,32235..32469,
32585..32716,32809..33174))
/gene="P0698A10.8"
/codon_start=1
/product="putative DNA polymerase I"
/protein_id="BAB92438.1"
/db_xref="GI:20804753"
/translation="WPPPAVSTAPVAPVAVSSSSAAAAAPRRPRSLRVRTPYASVSS
SSSPSTFAVISIPPPPAHRSKQLVRGDAAPAKPRVFLDVNPLFCRGSORSIGA
FARWALFAHVSLSDDPVAVLDGEGNEYRRLPSYKAHRPGRPGTGADSRVIDVL
RCNPFVVRVDGTEADDDVATUTEQVLQGRFVIGSPDKQFKQLISDVQVLPPIE
IGRWFSFYLRHVYAQKCDPTADLSLRCFIDEADGIPGIQHLVPGFRKTAVALKKK
HGSLENLNTAAVRTVGDKYAADALVKHADYLRKNYEVLRLKRDVKVQLDRLWLSRD
SCNDSVLSDFLTKNEGONIR"
complement(34177..34410)
/gene="P0698A10.9"
CDS
complement(34177..34410)
/gene="P0698A10.9"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB92439.1"
/db_xref="GI:20804754"
/translation="WTSUTGRAFDNRDDNGFIPATELARSMAADTNTDLHVTSPTPAARR
SWPPFSLLLTASRRMPGKKQKTKRRKKLEI"
complement(join(35312..35473,36485..36646))
/gene="P0698A10.10"
CDS
complement(join(35312..35473,36485..36646))
/gene="P0698A10.10"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB92440.1"
/db_xref="GI:20804755"
/translation="MEHGRGEPEGRGTGWVVDAPADCTCESGYVWAGPFFWAVST
NRLRLSDILQESGIPVPIQFLQOLQASKRLNLFSENRSTRDRFPKTPSDSKYSND
ETVRL"
join(37075..38148,39504..39906,40314..40415,40513..40736)
/gene="P0698A10.11"
CDS
join(37075..38148,39504..39906,40314..40415,40513..40736)
/gene="P0698A10.11"
/note="contains ESTs AU101336 (E11055), AU101337 (E11055)"
/codon_start=1

Query Match      8.0%; Score 115.6; DB 8; Length 137348;
Best Local Similarity 76.3%; Pred. No. 2.6e-11;
Matches 142; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 893 TTTCAAACGCATGATAACGAGAAAGCTCATTAGCACTATTACTTAGATATTATAT 952
|||||
Db 136817 TTTTAACTCAGCAAAACGAGAAACCTTATTAGCATAATTAATTAAGTATTATAAT 136758
|||||

QY 953 TATAAACTTGAAGAAAATATTTATTTGAATTTTAAACAATGTATGCATAATTTT 1012
|||||
Db 136757 TATAAACTTGAAGAAAATATTTATTTGAATTTTAAATAAATCTTTTATATAGATTTT 136698
|||||

QY 1013 TTAAAAACACACCAATTTAAACCCCTTTAAAAAGCATCTCTAATAGAAACGAGGAAGTAAA 1072
|||||
Db 136697 TTTAATTAACATATCATTTAAGTTTGAAGTGAAGTAACTACTACGGAACGATAAATAA 136638
|||||

QY 1073 GATTCA 1078
|||||
Db 136637 GTTTGA 136632
```

RESULT 4  
AP005594

```

LOCUS      AP005594                124815 bp      DNA      linear      HTG 24-JUL-2002
DEFINITION      Oryza sativa (japonica cultivar-group) chromosome 9 clone P0701E06,
                  *** SEQUENCING IN PROGRESS *** , in ordered pieces.
ACCESSION      AP005594
VERSION        AP005594.1  GI:21952963
KEYWORDS       HTG; HTGS PHASE2
SOURCE         Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
                  clone:P0701E06.
ORGANISM       Oryza sativa (japonica cultivar-group)
                  Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
                  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
                  Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE      1  Sasaki,T., Matsumoto,T., Hattori,M., Sakaki,Y. and Katayose,Y.
AUTHORS        Oryza sativa nipponbare(GA3) genomic DNA, chromosome 9, PAC
TITLE          clone:P0701E06
JOURNAL        Published Only in Database (2002)
REFERENCE      2  (bases 1 to 124815)
AUTHORS        Sasaki,T., Matsumoto,T., Hattori,M., Sakaki,Y. and Katayose,Y.
TITLE          Direct Submission
JOURNAL        Submitted (23-JUL-2002) Takuji Sasaki, National Institute of
                  Agrobiological Sciences, Rice Genome Research Program; Kannondai
                  2-1-2, Tsukuba, Ibaraki 305-8602, Japan
                  (E-mail:tsasakienias@affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
                  tel:81-298-38-7441, fax:81-298-38-7468)
COMMENT        NOTE: It currently consists of 1 contigs. Gaps between the contigs
                  are represented as runs of N. The order of the pieces is believed
                  to be correct as given, however the sizes of the gaps between them
                  are based on estimates that have provided by the submitter. This
                  sequence will be replaced by the finished sequence as soon as it is
                  available and the accession number will be preserved.
                  * NOTE: This is a 'working draft' sequence.
                  * This sequence will be replaced
                  * by the finished sequence as soon as it is available and
                  * the accession number will be preserved.
FEATURES             Location/Qualifiers
     source           1..124815
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /chromosome="9"
                     /clone="P0701E06"
BASE COUNT        37055 a 25766 c 25390 g 36503 t 101 others
ORIGIN
Query Match      7.6%; Score 109.2; DB 2; Length 124815;
Best Local Similarity 63.2%; Pred. No. 3.5e-10;
Matches 168; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 820 TAAGAACGATGAGTTGGGAAACAAAGAGAAAAACACAGCCCTTAAAGCCTTCTTGATCCTCTA 879
|||||
Db 56107 TTATAAGTATCTGTTAGGATTAAATCAATTAAAGATGACTAATGCTCGGTTCAATCACCTC 56166
|||||

QY 880 GTTGGAGTTGAATTTTCAACGCATGATAACGAGAAAGCTCATTAGCACATATTACTT 939
|||||
Db 56167 ATTTGGAGATAGAATTTTCAACGCACACAAACGAGAAATCTCATTAGCATGATCAAT 56226
|||||

QY 940 AGATATTTATATAATTATAAATTTGAAAAAATATTATTATTGAAATTTTAAACAATGTATG 999
|||||
Db 56227 AAGTGTTAATTAATTATAAATCTTGAAAAATGGATTTATTGTATATTATTAACACTTCTA 56286
|||||

QY 1000 CATAAATTTATTTTAAAAACACACCAATTTTAACCCCTTTAAAAAGCATCTCTTAATAGGAAA 1059
|||||
Db 56287 TATAGATTTTATTTTAAAGAAACACATCATTTTAACTGTTTGAAGAAAGTATGCTTAATGGA 56346
|||||

QY 1060 CGAGGAAGTTAAAGATTCCACGAAGT 1085
|||||
Db 56347 TCAGAGAAAGTTGAAGTTTACTAGAGT 56372
|||||

RESULT 5
AP003261/c
LOCUS      AP003261                138025 bp      DNA      linear      PLN 10-AUG-2002
```

DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, PAC clone:P0471B04.

ACCESSION AP0033261

VERSION AP0033261.3 GI:22202658

SOURCE Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA, clone:P0471B04.

ORGANISM Oryza sativa (japonica cultivar-group)

SpERMatoPhyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1

AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.

TITLE Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC clone:P0471B04

JOURNAL Published Only in Database (2001)

AUTHORS 2 (bases 1 to 138025)

TITLE Sasaki,T., Matsumoto,T. and Yamamoto,K.

JOURNAL Direct Submission

Submitted (19-FEB-2001) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan

(E-mail:tsasakienia@affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468)

COMMENT On Aug 9, 2002 this sequence version replaced gi:21202837. Genes were predicted from the integrated results of the following: GENSCAN1.0, BLAST2.0, BLASTX2.0 as well as SplicePredictor (October 1998 version). The genomic sequence was searched against NCBI Nonredundant Protein database, nr. (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP. Protein homologies of the coding regions were searched against NCBI Nonredundant Protein database with BLASTP2.0. ESTs represent the identified cDNA sequences using BLASTN 2.0 with the corresponding DBJ accession no. and RGP clone ID.

A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted with a gene prediction program is classified as a 'hypothetical' protein.

The orientation of the sequence is from SP6 to T7 of the PAC clone. This sequence of P0471B04 clone has an overlap with P0698A10 clone (DBJ : AP003297) at the position 1 to 24,438 of 5' end and an overlap with P0018C10 clone (DBJ: AP003227) at the position 48,801 to 138,025 of 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at http://rgp.dna.affrc.go.jp/Genomeseq.html.

Location/Qualifiers

1. 138025

/organism="Oryza sativa (japonica cultivar-group)"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/db\_xref="GI:22202658"

/chromosome="1"

/clone="P0471B04"

join(9533..9715,9887..10019,10850..10988,11365..11401,13640..13931,14823..14955,15564..15576)

/gene="P0471B04.1"

join(9533..9715,9887..10019,10850..10988,11365..11401,13640..13931,14823..14955,15564..15576)

/gene="P0471B04.1"

/note="hypothetical protein similar to Arabidopsis thaliana chromosome 1, F10C21\_14"

/codon\_start=1

/protein\_id="BAC07317.1"

/db\_xref="GI:22202659"

/translation="MSPFSVIGQFDTTYTKVFGGLAWETQKTMKRYFEQFGEILLE AVITDKTGRSGYGVTRDPDPAAMRACVDPAVIDGRRANCLASLGQVRSRPT PQHCARSFRVMKSFQOAGIQGGIAAFPSHAFPHYAI PQGLPYHYGVSPYSPDY SPYNNYNIYVGGADY PFGVAAAAAAGAAAGWVTGSSPFYVYFQFGSGSTTTNVASQ GYNLYQPMHFSTVASTAAVTGAQYGGPLSLAASPAQAQAAKANGKSNPNQAL TQGGFMUQANWTKVTSKLTIEQCHFKI5EKP"

complement (join(17199..17342,17914..18008,18141..18178,

18784..19073))

/gene="P0471B04.2"

complement (join(17199..17342,17914..18008,18141..18178,18784..19073))

/gene="P0471B04.2"

/codon\_start=1

/product="putative SnRK1-interacting protein 1"

/protein\_id="BAC07318.1"

/db\_xref="GI:22202660"

/translation="MASGALARLLGRRAAATPLLARPPFAAKASRRPOEPAPFSED EDPFAGEVAAPAPTEGISKPLAEVLRELCKRVPALVKTRVEDGFSLKVIPWHIVNK LNIHAPFGEVRSIVISDGSKSVIYRTVLGTDAEIYREATGTSADDTGYGDP VQKAEMAFRRACRTRIGLGLUHLHEDMS"

join(24593..24925,25029..25403)

/gene="P0471B04.3"

join(24593..24925,25029..25403)

/gene="P0471B04.3"

/note="hypothetical protein"

/codon\_start=1

/protein\_id="BAC07319.1"

/db\_xref="GI:22202661"

/translation="MARRGGGVILRGRCRRRGERGEQVRPLPLASTVTGGRGGR GRRRGRERAGCTARQSSSPRSSGRRSSSNWTRSEHQIILSTDOSRTCLSNWDS TGLPVSSKVIRORGAGVRRRGRSGRRGRGAGAAVQDGGWRGCGHGRGRSRGA GMAANDGADAAAADADGGLGGCGRSPFIRRAALLWQSSSSRRRFPFPTPLAAPS PSAVLLSSGAMLPNRR"

join(25836..26033,26372..26471,26764..26924,27200..27397,27476..27568,27766..27888,28671..28918,29038..29209,29441..29527,29557..29820,30449..30611,30873..30960,31044..31113,31600..31728,32356..32445)

/gene="P0471B04.4"

join(25836..26033,26372..26471,26764..26924,27200..27397,27476..27568,27766..27888,28671..28918,29038..29209,29441..29527,29557..29820,30449..30611,30873..30960,31044..31113,31600..31728,32356..32445)

/gene="P0471B04.4"

/note="hypothetical protein similar to Arabidopsis thaliana chromosome 5, At5g49880"

/codon\_start=1

/protein\_id="BAC07320.1"

/db\_xref="GI:22202662"

/translation="MILRTPQKRRAADSDVDAATAATTRSPVSDRLVLYDRPTAL VPAGPGEPMDDMVCTYHCRQVKSFPVALDTAEKQVQBYRATKLDDBEKLKSEDE RACQDKNLNVEQELAAATKGRSAQMYERLLKEVDFQERYCDQIKIIGELETQKKEI DSRMAESSASAKESVKELEGNLISENREKALKKLSYLODDTKLSIKLNA ELBRMLRALNSDEAKLLEQDLKQLDSESVRENEMHRLNCSLSYERTPSD DQKLLQLEELANYKEVDEARRLSSHTNVLLKEKILEEQCERAEMLSKLOE IEAKQKLELASCATALLSNIPDVSSFGDIPQKIADLOQALNLNKNVETVSQLKE LKVALEFADLUSQKRAEATLAKERASATREIKRLELLAAASEDRDLRKDAVSK SRGDDASSKFCAPILNGMHNISCEVTIQNMESDLSRMEKVVTELESTIRDRLELIS QHTELNLMNEKLSIESRKAKSLERDQDLRSOVALLESKLGHDYASSTKVLRMVN TLAVDNEAKOTIEALQAEKTKERLQAEVBEKGADVTVDVNIAEKLAQLKNQIA TLEKREERYKVAFAERISVFRKACCSLFGYKIVMDDQSGNGIPVTRFILHSVVAQSD DEKLEDFESGSGTNIIVDFIRKMSIPATAULTMESFNKRSIC"

join(33470..33475,33573..33794,33969..34033,34129..34322,34573..34839,34942..35029,35125..35250,35325..35434,35557..35705)

/gene="P0471B04.5"

join(33470..33475,33573..33794,33969..34033,34129..34322,34573..34839,34942..35029,35125..35250,35325..35434,35557..35705)

/gene="P0471B04.5"

/note="contains ESTs C91865 (E31978), AU166151 (E31978)"

/codon\_start=1

/product="putative elicitor response protein"

/protein\_id="BAC07321.1"

/db\_xref="GI:22202663"

/translation="MMTLSTSRQPEQKPPRTPPLPGKVAALCVASFVGGLLLS GRVVVPLLPGGSSPNASYKTSFTGCENKRAKLGSDTINNEVSRTHHAIASLDK AVSLEMELAEVARSAAVAGTASGLQKAFVVGINTAFSSKRRDSLRDTWV PRGDKRLRLKEKGIIVRFVIGSAGAAAGDPLDRAVDAEAKENGLFLRDJHVEGYH ELSKTRVITYTTAVATWADFPYKVDVDDVHVNGLMGLTSRLAKYTRPRVYGVCKSGP VLSQKGVKYHEPBYWKFDEGKNKFRHATCQITAYVSKDLAAYISINQILHRFANEDV

FEATURES

source

gene

CDS

gene

```
SILGAWLIGLEVEHVDRLSMDCCATPPDCWKKRAGNVCVASFDMSCSGVCKSVDRMKHI
HRACGEGQAVWSVAT"
join(36438..36570,36637..36956)
/gene="P0471B04.6"
CDS
join(36438..36570,36637..36956)
/gene="P0471B04.6"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC07322.1"
/db_xref="GI:22202664"
/translation="MSLLTPSLVVKPSPRCRPOYRGRGSMSSVCGPGREHHWG
PAERERREYRRRRGRRRVILPNNIGACPLSPRCGADGTATRRSRGGPDAP
TTAPLPLPVTVRDGEVSDGARQWRQWRDQRGREREATAAD"
complement(join(38223..38304,39187..39287))
/gene="P0471B04.7"
complement(join(38223..38304,39187..39287))
/gene="P0471B04.7"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC07323.1"
/db_xref="GI:22202665"
/translation="WVKNDRSVTGLTPVTCWDNNAPTSHLLPAEWSTFHSIDFT
CKVIKRIIVKSYSAIG"
join(43057..43393,43868..43997,44238..44453,44554..45171,
45260..46109)
/gene="P0471B04.8"
join(43057..43393,43868..43997,44238..44453,44554..45171,
45260..46109)
/gene="P0471B04.8"
/note="contains ESTs AU057294 (S21318), AU057295 (S21318)"
/codon_start=1
/product="putative male sterility 1 protein"
/protein_id="BAC07324.1"
/db_xref="GI:22202666"
/translation="WVNSRPLKRAITRVEARDFAFPAGDGGAGTFRRAVGFLLA
RYALPLPSPVSPAAAPPHLLTRVSLRVGDEEGGGAVELNVVEEDVLRSR
SYVQDQRPVSGSHVPCGRYHFIENDNNQVCGKRSCCLRCGCTPTVAGESRCLLC
NFDPMGELEECGYMHLDDNTHLLHAVHANGYHLLRVNGREGGSRCLTGRDIMSEF
DLCKVLHVKTVMVDISKHGMVYRLHAIITSGHPWYGEWYKFGAGSFALTSDTYQ
EAVDTLSGIALYKFSHPQIETPLQNTIALYWALSRLQVTRDLPFPIHLLHQR
KNETSKEPTIDHKSVASNLCKTEDIIDRAETAMKLVRVQPGQWVRALRGAR
SKAVDSQELLDYSLKGLGKMDGHHFIAVRNAAETSAIEYRLDENSNSQSDVDAAFGP
SYDHLHLKFLYNALNPETMLASQPEVIGASSHAAAKTIDCKQFIKHVDQHTPRA
PLNPFLLSVRCISIELLDHPKQYTAPELVLLIPASATIAELKIQATRAFOETLYMFO
YVEOLPDPNPSIDTLVKHVLGSSQLVRGRCTGDNRRIVQFRMERGLENNWTVDCT
CCAKDDGGERMLACDVGWQHTRCSGISDFDVPKEFICRKCASPRGKRGGGGGN
GGSRDVDAAGRCKDEIGSSVGGAGKFRMATVG"
join(54852..54959,55057..55188,55283..55412,55519..55605,
56032..56102)
/gene="P0471B04.9"
join(54852..54959,55057..55188,55283..55412,55519..55605,
56032..56102)
/gene="P0471B04.9"
Query Match 7.6%; Score 109.2; DB 8; Length 138025;
Best Local Similarity 76.3%; Pred. No. 3.5e-10;
Matches 142; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 893 TTTCACACGATCAAAACGAGAACTCATAGACATATTACTTAGATATTATAT 952
Db 23907 TTTTAACTCACCAAAACGAGAAACTTATTAGACATAATTAATTAAGTATTATAT 23848
QY 953 TATAAACTTGAAAAAATATTTATTGAAATTTTAAACAAATGATGCAATAATTTT 1012
Db 23847 TATAAACTTGAAAAAATATTTATTGATTTTAAAAATAACTTTTATATAGATTTT 23788
QY 1013 TTAATAACACACCAATTTAACCCCTTTAAAAAGCATCCTTAATAGGAACGAGGAAGTTAA 1072
Db 23787 TTAATAACACATATCAITTTAACAGTTTGAAGAGCATACTACGAGAAACGATTAATTA 23728
QY 1073 GATTCA 1078
Db 23727 GTTTGA 23722
```

```
RESULT 6
CNS07YPX
LOCUS
DEFINITION
CNS07YPX 169663 bp DNA linear HTG 27-JUL-2002
Oryza sativa chromosome 12 clone OUL1618_C05, *** SEQUENCING IN
PROGRESS ***, in ordered pieces.
ACCESSION
AL713952
VERSION
AL713952.2 GI:22003292
KEYWORDS
HTG; HTGS PHASE2; HTGS_ACTIVEFIN.
SOURCE
Oryza sativa.
ORGANISM
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 169663)
Choisne,N., Orjeda,G., Cattolico,L., Demange,N., Wincker,P.,
Segurens,B., Pelletier,E., Scarpelli,C., Salanoubat,M.,
Weissenbach,J. and Quetier,F.
Oryza sativa chromosome 12 sequencing
Unpublished
2 (bases 1 to 169663)
Genoscope.
Direct Submission
Submitted (26-JUL-2002) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
COMMENT
On Jul 29, 2002 this sequence version replaced gi:19715885.
IMPORTANT: This sequence is unfinished and does not necessarily
represent the correct sequence.
Work on the sequence is in progress and the release of this data is
based on the understanding that the sequence may change as work
continue. The sequence may be contaminated with foreign sequence
from E.coli, yeast, vector, phage, etc.
The nucleotide sequence of this BAC clone was generated by
combining Monsanto and Genoscope sequencing data.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
source
1.169663
/organism="Oryza sativa"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:4530"
/chromosome="12"
/clone="OUL1618_C05"
/clone_lib="Monsanto"
BASE COUNT 46256 a 37378 c 36384 g 49645 t
ORIGIN
Query Match 7.6%; Score 109.2; DB 2; Length 169663;
Best Local Similarity 68.8%; Pred. No. 3.3e-10;
Matches 150; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 861 AAGCTCTCTTGATCCTCTAGTTCGAGGCTTGATTTTCAACGCATGATAACGAGAAAGCT 920
Db 103185 AAGCCTAGGTTTCGGCTTGTTGGATAGAAAATTTTAAACGCAGCAAAAGAAAGCT 103244
QY 921 CATTAGCACATTTACTCTAGATATTTATATATAAATTTGAAAAAATATTATTTTGA 980
Db 103245 CATTAGCACATAATTGATTAAAGTATTAAATATATAAATTTGAAAAATGATTTATTGA 103304
QY 981 ATTTTAAACAATGATGATGATTAATTTTAAAAACACACCAATTTTAAACCCCTTTAA 1040
Db 103305 TTTTAAACAACCTCTATATAGAAATTTTCTTAAAAAACAATCATCTTTAAACAGTTTAA 103364
QY 1041 AAAGCATCTTAATAGGAACGAGGAAGTTAAAGATTCA 1078
Db 103365 AAAGCGTGCTAACAGAAACGAGGAAGTTGAAGTTTGA 103402
```



```

RESULT 7
LOCUS OSJN00015 126659 bp DNA linear PIN 17-JUL-2002
DEFINITION Oryza sativa genomic DNA, chromosome 4, BAC clone: OSJNB00079B02, complete sequence.
ACCESSION AL606455
VERSION AL606455.2 GI:21740404
KEYWORDS HTG
SOURCE Oryza sativa.
ORGANISM Oryza sativa
REFERENCE 1
AUTHORS Han, B., Feng, Q., Huang, Y.C., Li, Y., Zhu, J.J., Zhao, Q., Hu, X., Liu, Y.L., Mu, J., Yu, Z., Chen, L., Fan, D.L., Weng, Q.J., Zhang, L., Lu, Y.Q., Yu, S.L., Liu, X.H., Lu, T.T., Zhang, Y.J., Lu, Y., Li, C., Li, T., Zhang, Y., Hu, H., Jia, P.X., Qian, Y.M., Ying, K., Zhou, B., Chen, Z.H., Hao, P., Zhang, L., Wu, M., Zhang, R.Q., Guan, J.P., Fu, G., Wang, S.Y., Ren, S.X., Lv, G., Lin, W., Gu, W.Q., Zhu, G.F., Tu, Y.F., Jia, J., Yin, H.F., Zhang, Y., Cai, Z., Chen, J., Kang, H., Chen, X.Y., Shao, C.Y., Sun, Y., Hu, Q.P., Zhang, X.L., Zhang, W., Wang, L.J., Ding, C.W., Sheng, H.H., Gu, J.L., Chen, S.T., Ni, L., Zhu, F.H. and Hong, G.F.
TITLE Direct Submission
JOURNAL Submitted (27-JUN-2002) Han Bin, National Center for Gene Research, Chinese Academy of sciences, 500# Cao Bao Road, Shanghai 200233, CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests: bhan@ncgr.ac.cn
REMARK Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC clone: OSJNB00079B02.
COMMENT On Jul 12, 2002 this sequence version replaced gi:15552693.
Web site: http://www.ncgr.ac.cn
----- Summary Statistics
Assembly program: phrap
-----
NOTE: This is a complete sequence.
Genes were identified by a combination of several methods: Gene prediction programs including Fgenesh (http://www.softberry.com/), genscan (http://CCR-081.mit.edu/GENSCAN.html), GeneMarkHM (http://genemark.biology.gatech.edu/Genemark/), tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/), searches of the complete sequence against NCBI none redundant protein database (nr) (ftp://ncbi.nlm.nih.gov/blast/db) and the EST database at NCGR.
FEATURES
source
1..126659
/organism="Oryza sativa"
/variety="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:4530"
/chromosome="4"
/clone="OSJNB00079B02"
/clone_lib="CUGI-OSJNB"
BASE COUNT 35469 a 27283 c 28042 g 35865 t
ORIGIN
Query Match 7.6%; Score 108.8; DB 8; Length 126659;
Best Local Similarity 55.9%; Pred. NO. 4.1e-10;
Matches 227; Conservative 0; Mismatches 177; Indels 2; Gaps 1;
QY 660 TAGACGGCTAACCCACCACCTTAAGCTTAAGTTCGAGAGGTGAAGTACGCACGCAAAAT 719
DB 15094 TAGATTGAACACCTTTCTTCCTTAGCACGTAAACGAAACAACATATTAGCCGCAATGA 15153
QY 720 ATGATGGTTTATTATATGATTTTTTTTAAATAAAGTTTACACATAAATTTCTTTAGGAAC 779
DB 15154 TTAGCAATATATTATATATATTATTAGCAACTTTCTTTTATTATTATTTTCATAAAGAC 15213
QY 780 ATATCATTTAATGGTTTGAACACGTGCAATAGAAACTAAGAACGATGAGTTGGGAA 839
DB 15214 GCATCATTTTAGTCGTTTGAACACGTGCAACGTGCAACCAATAGACTGAAGTTAAAAA 15273
QY 840 ACAAGAGAAAAACACACGCTTAAGGCTTCTTGATCCTCTAGTTGGAGGTTGATTTTCAA 899

```

---

```

Db 15274 ATTGGAAGAAATATACTAAGGCCCGTTCGATCTCCAGAGTGGAGATAGATTTTAT 15333
QY 900 CGCATGATAACGAGAAAGCTCATTAGCACATTTACTTAGATATTATTATTAATAAAC 959
Db 15334 GGCAGAAAAACGAGAAATCTCTTAACACATGATTAATAAATATTAACTATTATAAAT 15393
QY 960 TTGAAAAAATATTATTGTAATTTTTTAAACAA--TGTATGCATAAATATTATTTTAAA 1017
Db 15394 TTGAAAAATGATTTATTGTGCTTTTAAACAACACTTGTTATATAGAAACTTGCTGTA 15453
QY 1018 AACACACCAATTTAACCTTTTAAAGAGCATCTCTAATAGGAACGAG 1063
Db 15454 TAATTTACCTTTTAACCTTTTGAAGAAACGCTGCTGATGGAATAAG 15499
RESULT 8
LOCUS OSJN00094/c
DEFINITION Oryza sativa chromosome 4 clone OSJNB0003B01, *** SEQUENCING IN PROGRESS ***, 3 ordered pieces.
ACCESSION AL606649
VERSION AL606649.2 GI:21912479
KEYWORDS HTG; HTGS PHASE2.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
REFERENCE 1
AUTHORS Han, B., Feng, Q., Huang, Y.C., Li, Y., Zhu, J.J., Zhao, Q., Hu, X., Liu, Y.L., Mu, J., Yu, Z., Chen, L., Fan, D.L., Weng, Q.J., Zhang, L., Lu, Y.Q., Yu, S.L., Liu, X.H., Lu, T.T., Zhang, Y.J., Lu, Y., Li, C., Li, T., Zhang, Y., Hu, H., Jia, P.X., Qian, Y.M., Ying, K., Zhou, B., Chen, Z.H., Hao, P., Zhang, L., Wu, M., Zhang, R.Q., Guan, J.P., Fu, G., Wang, S.Y., Ren, S.X., Lv, G., Lin, W., Gu, W.Q., Zhu, G.F., Tu, Y.F., Jia, J., Yin, H.F., Zhang, Y., Cai, Z., Chen, J., Kang, H., Chen, X.Y., Shao, C.Y., Sun, Y., Hu, Q.P., Zhang, X.L., Zhang, W., Wang, L.J., Ding, C.W., Sheng, H.H., Gu, J.L., Chen, S.T., Ni, L., Zhu, F.H. and Hong, G.F.
TITLE Direct Submission
JOURNAL Submitted (27-JUN-2002) Han Bin, National Center for Gene Research, Chinese Academy of sciences, 500# Cao Bao Road, Shanghai 200233, CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests: bhan@ncgr.ac.cn
REMARK Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC clone: OSJNB0003B01.
COMMENT On Jul 19, 2002 this sequence version replaced gi:15594106.
Web site: http://www.ncgr.ac.cn
----- Summary Statistics
Assembly program: phrap
-----
NOTE: This is a PHASE2 sequence. Gaps are shown by nnnn. Genes were identified by a combination of several methods: Gene prediction programs including Fgenesh (http://www.softberry.com/), genscan (http://CCR-081.mit.edu/GENSCAN.html), GeneMarkHM (http://genemark.biology.gatech.edu/Genemark/), tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/), searches of the complete sequence against NCBI none redundant protein database (nr) (ftp://ncbi.nlm.nih.gov/blast/db) and the EST database at NCGR.
* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 61924: contig of 61924 bp in length
* 61925 62124: gap of 200 bp
* 62125 76571: contig of 14447 bp in length
* 76572 76771: gap of 200 bp

```



VERSION	D10134.1	GI:391876
KEYWORDS	adenylate kinase-a; Adk-a.	
SOURCE	Oryza sativa (strain: Famanoushi) callus cDNA to mRNA, clone lib1ambda gt10 clone: Adk-a-29 and Adk-a-50.	
ORGANISM	Oryza sativa	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.	
REFERENCE	1 (bases 1 to 1154)	
REFERENCE	2 (bases 1 to 1154)	
TITLE	Molecular characterization of cDNA encoding for adenylate kinase of rice ( <i>Oryza sativa</i> L.)	
JOURNAL	Plant J.	2 (6), 845-854 (1992)
MEDLINE	93258424	
REFERENCE	2 (bases 1 to 1154)	



```

* 76772 155582: contig of 78811 bp in length.
FEATURES
  source
    Location/Qualifiers
      1..155582
        /organism="Oryza sativa"
        /variety="Nipponbare"
        /sub_species="japonica"
        /db_xref="taxon:4530"
        /chromosomes="4"
        /clone="OSJNB0003B01"
        /clone_lib="CUGI-OSJNB"
BASE COUNT 40185 a 36681 c 38162 g 40148 t 406 others
ORIGIN
  Query Match 7.6%; Score 108.8; DB 2; Length 155582;
  Best Local Similarity 55.9%; Pred. No. 4e-10;
  Matches 227; Conservative 0; Mismatches 177; Indels 2; Gaps 1;
Qy 660 TAGACGGCTAAGCCACCCATCTAAGCTAAGTTCGAGAGGTGAAGTACGACGAAAT 719
Db 14837 TAGATTGAGAACCTTTCTCTTAGCACGTAAAAACGAAACACATATTAGCGCATAA 14778
Qy 720 ATGATGGTTTAAATATGATTTTTTTTAAATAACTTTCACATAAATTTCTTTAGGAAC 779
Db 14777 TTAGCAATATTAATAATAATTTTGAAGCAACTTCTTTTAAATTTTATTTCATAAAGAC 14718
Qy 780 ATATCATTTAATGGTTTGAACAACTGCACATAGAAACCTAAGAACGATGAGTTGGGAA 839
Db 14717 GCATCATTTTGTGCGTTTGGAACTGCAGCTGAAACCAATAGACTAGAAAT 14658
Qy 840 ACAAGAGAAAAACACAGCCTTAAGGCTTCTTGATCCTCTAGTTGGAGGTGATTTTCAA 899
Db 14657 ATTGGAAGAAATAATACTAAGGCCCGTTCGATCTCCAGGAGTGAGATTTTAT 14598
Qy 900 CGCATGATTAACGGAAGCTCATAGCACATTTATTTAGTATTTTATATATATATAC 959
Db 14597 GGCAGAAACAGAGAAATCTCATTAACATGATTAAATTAATAATTAATTTTAAAT 14538
Qy 960 TTGAAAAAATATTTATTTGAAATTTTTTAAACAA--TGTATGCATAAATTTATTTTAA 1017
Db 14537 TTGAAAAATTTGATTTATTTCTTTTAAACAACTTGATATAGAACTTGTTGTAATA 14478
Qy 1018 AACACCAATTTAAACCTTTAAAGCATCTTAATAGGAAACGAG 1063
Db 14477 TAATTTACCTTTTAACCTGTTTGAACAACTGCTCATGGAATAAAG 14432

RESULT 9
OSIG00047 100419 bp DNA linear HTG 01-AUG-2002
LOCUS Oryza sativa chromosome 4 clone H0322F07, *** SEQUENCING IN
DEFINITION PROGRESS ***, in ordered pieces.
ACCESSION AL732355
VERSION AL732355.1 GI:22091860
KEYWORDS HTG; HTGS PHASE2.
SOURCE Oryza sativa
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzae; Oryza.
REFERENCE
  1 Han, B., Chen, Z.H., Feng, Q., Huang, Y.C., Li, Y., Zhu, J.J., Zhao, Q.,
    Hu, X., Liu, Y.L., Mu, J., Yu, Z., Chen, L., Fan, D.L., Weng, Q.J.,
    Zhang, L., Lu, Y.Q., Yu, S.L., Liu, X.H., Lu, T.T., Zhang, Y.J., Lu, Y.,
    Li, C., Li, T., Zhang, Y., Hu, H., Jia, P.X., Qian, Y.M., Ying, K.,
    Zhou, B., Hao, P., Zhang, L., Wu, M., Zhang, R.Q., Guan, J.P. and
    Hong, G.F.
    Direct Submission
    Submitted (05-MAY-2002) Han Bin, National Center for Gene Research,
    Chinese Academy of sciences, 500# Cao Bao Road, Shanghai 200233,
    CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests:
    bhan@ncgr.ac.cn
    Oryza sativa indica (guangluai4) genomic DNA, chromosome 4, BAC
    clone: H0322F07.

```

```

Web site: http://www.ncgr.ac.cn
----- Summary Statistics
Assembly program: phrap
-----
This is a complete sequence.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
  source
    Location/Qualifiers
      1..100419
        /organism="Oryza sativa"
        /variety="guangluai4"
        /sub_species="indica"
        /db_xref="taxon:4530"
        /chromosomes="4"
        /clone="H0322F07"
        /clone_lib="BAC-HindIII"
BASE COUNT 29642 a 21538 c 21324 g 27915 t
ORIGIN
  Query Match 7.5%; Score 108.4; DB 2; Length 100419;
  Best Local Similarity 72.9%; Pred. No. 5.1e-10;
  Matches 153; Conservative 0; Mismatches 56; Indels 1; Gaps 1;
Qy 869 TTGATCCTCTAGTTGGAGTTGATTTTCAAAACGATGATAAAGCAGAAAGCTCATTAGCA 928
Db 51985 TTGATCATCGTTTGGCAGATAGATTTTGTCTACACGCAAAACGAGAAATCTCATTAGCA 52044
Qy 929 CATTTACTTTAGATATTTTATAATTAATAACTTTGAAAAAATATTTATTTTGAATTTTTT 988
Db 52045 CATGATTAATTAAGTATTAAGTATTAATAATTTTAAAAATGAATTTATTTTG-CTTTTTT 52103
Qy 989 AACATGTATGCATATTTATTTTAAAAACACACCAATTTAACCTTTAAAAAGCATC 1048
Db 52104 AACACTTGTATTAATAACTTTTAAAAACGTAACCATTTAACGTTTGAAGAGCGTG 52163
Qy 1049 CTAATAGAAACGAGAGTGTAAAGATTCA 1078
Db 52164 CTAATGAAACGAGAAAGTTGAAGTTTGA 52193

RESULT 10
AP003446 100635 bp DNA linear PLN 23-MAY-2002
LOCUS Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
DEFINITION BAC clone: OJL529_G03.
ACCESSION AP003446
VERSION AP003446.3 GI:21104777
KEYWORDS Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
SOURCE clone: OJL529_G03.
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzae; Oryza.
REFERENCE
  1 Sasaki, T., Matsumoto, T. and Yamamoto, K.
    Oryza sativa nipponbare (CA3) genomic DNA, chromosome 1, BAC
    clone: OJL529_G03
    Published Only in Database (2001)
  2 (bases 1 to 100635)
    Sasaki, T., Matsumoto, T. and Yamamoto, K.
    Direct Submission
    Submitted (28-MAR-2001) Takuji Sasaki, National Institute of
    Agrobiological Sciences, Rice Genome Research Program; Kannondai
    2-1-2, Tsukuba, Ibaraki 305-8602, Japan
    (e-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
    Tel:81-298-38-7441, Fax:81-298-38-7468)
    On May 22, 2002 this sequence version replaced gi:16041055.
    Genes were predicted from the integrated results of the following:
    GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor
    (October 1998 version). The genomic sequence was searched against

```

NCBI NonRedundant Protein database, nr  
(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at  
RGP. Protein homologies of the coding regions were searched against  
NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent  
the identified cDNA sequences using BLASTN 2.0 with the  
corresponding DDBJ accession no. and RGP clone ID.

A gene with identity or significant homology to a protein is  
classified based on the protein name to indicate the homology level  
such as same name, 'putative-' and '-like protein'. A gene without  
significant homology to any protein but with EST homology (covering  
almost the entire length of partial sequence) is classified as an  
'unknown' protein. A gene predicted with a gene prediction program  
is classified as a 'hypothetical' protein.

The orientation of the sequence is from -21m13 to M13rev of the BAC  
clone. This sequence of OJ1529 G03 clone has an overlap with  
P0454H12 clone (DDBJ : AP003255) at the position 1 to 31,196 of 5'  
end and an overlap with P0703B11 clone (DDBJ : AP003302) at the  
position 85,137 to 100,635 of 3' end. Detailed information on  
overlap and assembly quality together with annotation of this entry  
is available at <http://rgp.dna.affrc.go.jp/GenomeSeq.html>.

## FEATURES

## source

```
1. 100635
/organism="Oryza sativa (japonica cultivar-group)"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="1"
/clone="OJ1529 G03"
join(28..136,803..1007,2468..2542,2784..3209,3357..3502,
4325..4443)
/gene="OJ1529_G03.1"
join(28..136,803..1007,2468..2542,2784..3209,3357..3502,
4325..4443)
/gene="OJ1529_G03.1"
/feature="hypothetical protein"
/codon_start=1
/protein_id="BAB93365.1"
/db_xref="GI:21104778"
/translation="MLRPVDNVHTEKRAQKLDCTDTEYISLSVMWIFFSPERRTRAHH
FALGRKLIVQNNLKEKERKPEPLAGAKTIIQTITAEGSLNKLKGNKPEPLKKE
GEDIQTVRHKSQVHKVODFELLVVNVKKNKKELTLETNPQTNLTILEESKKRTC
NOTQTDPLQSEQRFAAPRSRGSRRGARKEIALTNELGKRERAPHLHPQRPQR
RGYTVTAATFQVQLSHSPNPIRTPLFLDRGSYDEGEVEAKLWRTSED
RVFTIGTRAKRGSGEESRAGKSAVQPSATAFACSGRTESVFDTIQVKNKRC
FSNNLWDFTKNTYDFIPNNLGS"
complement(join(5387..5424,5515..5711,6719..7027,
7922..8039,8200..8246,9341..9463,11128..11291))
/gene="OJ1529_G03.2"
complement(join(5387..5424,5515..5711,6719..7027,
7922..8039,8200..8246,9341..9463,11128..11291))
/gene="OJ1529_G03.2"
/feature="hypothetical protein"
/codon_start=1
/protein_id="BAB93366.1"
/db_xref="GI:21104779"
/translation="MPARQFEDTNNRKPPELDTLLLLPSLPVPEAARRRGWE
SPAARDPRREYVHQRIWLDNRMAACPRLNRSEVGRSARIRIARTDARDKHGYK
YSTKADWAPLKLHVLKHYSYVNCVEVIGLECSWIKVSSVLESRRPLSLRV
RSRRLAGALSPPLAAGCCDYGDLHLAASNSDLGNKRRRAAABODDLKRLRV
HAGMGQCNLPFSFSSFLNFCVFLIETHRAILERKIIAYLTTRGEEKDRCPK
STYTKGVCTFYRRCENACISQTRVLFDFPKVTKHNSNLMVAESQVKGIANL"
join(13718..14497,14615..15031)
/gene="OJ1529_G03.3"
join(13718..14497,14615..15031)
/gene="OJ1529_G03.3"
/feature="hypothetical protein"
similar to Arabidopsis thaliana chromosome 3, F17014.14"
/codon_start=1
/protein_id="BAB93367.1"
/db_xref="GI:21104780"
/translation="WMSSTIAKTTSQAQNSRPSPTPTSPRIMSSSSSTSAAGRPS
SGRTSTTSRSTSSSTSPVSSRSSRSTPTQPIRTSSAPLAGHSVGRIFGNNI
TICRPTVTSNGRSPASSRSPSSPNSRARAPVRPLDIPDPSPTPNLRKTLPOR
PLSAGRARPVGGLGPKAPNAQVRSAPVKMTVPAITRSKFPDAPSRVSLTNGHOS
```

gene

CDS

gene

CDS

gene

CDS

```
gene      join(40305..40374,40975..41102,41389..41493,41889..42032)
          /gene="OJ1529.G03.7"
CDS       join(40305..40374,40975..41102,41389..41493,41889..42032)
          /gene="OJ1529.G03.7"
          /codon_start=1
          /product="ubiquitin conjugating enzyme"
          /protein_id="BAB93371.1"
          /db_xref="GI:21104784"
          /translation="MASKRILKELKDLQKDPPTSCSAGPAGEDMFHQATIMGPPDSP
YAGGVFLNHFPPDPYFKPKVSKFKVPHFNINSKSLCLDLKEQWSPALTIISKV
LLSICSLTDPNPDPDLVPEIAHMYKTRDKPETYTARSQKYANG"
          46484..47284
gene      /gene="OJ1529.G03.8"
          46484..47284
CDS       /gene="OJ1529.G03.8"
          /codon_start=1
          /product="putative gypsy-type retrotransposon RIRE8B ORF0
protein"
          /protein_id="BAB93372.1"
          /db_xref="GI:21104785"
          /translation="MPRYKPEYLFQVGFQELRTWMSFVIGFNAPPYTOIPIHYHEE
KCRVKVILNSSEVPFAGGRNVIHACQEVARIAGELDRYSDQLADREYEVH
PRQPGSDRGSYLGTENDAITRHILVEMWDESRATVLAQDREDRNGKICK
LEDKVDRLKELAAKGEAPQKARVRLTARKRALFVRYQLAPKRVVKEVAPTIA
DPPVIIISDDEEGSKRTRSKIEWVATQDDEDFNEPSIDLRCSELE"
Query Match      7.5%; Score 107.6; DB 8; Length 100635;
Best Local Similarity 62.0%; Pred. No. 7.1e-10; Mismatches 104; Indels 0; Gaps 0;
Matches 170; Conservative 0;

QY 853 ACAGCCTTAAGCCTTCTGATCCTCTAGTTGGAGGTTGATTTTCAAACGCATGATAAACG 912
Db 89942 ATAACTTAAGCCCTTCGATCTCTGAGTGAGACAGATTTTCAGCGCAGCAAAACG 90001

QY 913 AGAAGCTCATAGACATTTACTTACATTTATTTATTTATTAATAACTTGAAAAATAT 972
Db 90002 AGAAGCTCATAGACATTTATTAAGTATTAATTTATTTATTAATAACTTGAAAAATGGT 90061

QY 973 TTATTGTAATTTTTTAAACAATGTATGCATAAATTTATTTTAAAAACACCAATTTAA 1032
Db 90062 TTATTGTAATTTCTAGACAACTTCTAGTAGAACATTTTGTCAAAAAGCACCATTAA 90121

QY 1033 CCCTTTAAAAAGCATCTTAATAGGAAACAGGAAAGTTTAAAGATTCCACGAAAGTTTGGGA 1092
Db 90122 TAATTTAAAAAGCATGTTTAAACGAAACAGGAGTTGAGAAAAAGAACTGGCGCTTAGTC 90181

QY 1093 TAATGAAATCGGGTGGGATAGATTGGTAAA 1126
Db 90182 AAACTTTGCTAAGTTTGACTTAAAAAATAGTCAA 90215

RESULT 11
LOCUS     AP003302
DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
PAC clone: P0703B11.
ACCESSION AP003302
VERSION   AP003302.2 GI:17026069
KEYWORDS  Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
SOURCE    clone:P0703B11.
ORGANISM  Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1
AUTHORS   Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE     Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
clone: P0703B11
JOURNAL   Published Only in Database (2001)
REFERENCE 2 (bases 1 to 144074)
AUTHORS   Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE     Direct Submission
```

```
JOURNAL   Submitted (19-FEB-2001) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program, Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,
Tel: 81-298-38-7441, Fax: 81-298-38-7468)
COMMENT   On Nov 20, 2001 this sequence version replaced gi:13027332.
Genes were predicted from the integrated results of the following:
GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor
(October 1998 version). The genomic sequence was searched against
NCBI NonRedundant Protein database, nr
(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at
RGP. Protein homologies of the coding regions were searched against
NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent
the identified cDNA sequences using BLASTN 2.0 with the
corresponding DBJ accession no. and RGP clone ID.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with EST homology (covering
almost the entire length of partial sequence) is classified as an
'unknown' protein. A gene predicted with a gene prediction program
is classified as a 'hypothetical' protein.
The orientation of the sequence is from SP6 to T7 of the PAC clone.
This sequence of P0703B11 clone has an overlap with OJ1529_G03
(DDBJ: AP003346) at 5' end and with P0485B12(DDBJ: AP003348) at 3'
end. Detailed information on overlap and assembly quality together
with annotation of this entry is available at
http://rgp.dna.affrc.go.jp/GenomeSeq.html.
FEATURES             Location/Qualifiers
     source            1..144074
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /chromosome="1"
     gene              /clone="P0703B11"
                     complement(1630..2052)
     CDS                /gene="P0703B11.1"
                     complement(1630..2052)
                     /gene="P0703B11.1"
                     /note="hypothetical protein"
                     /codon_start=1
                     /protein_id="BAB85283.1"
                     /db_xref="GI:18844813"
                     /translation="MAVEARTSIGREEPTVCLPSPTRPPRLNHVAGKPNRCRF
LQIRROGHSTTKVGRRAAGWDDEERWEEETGLPGAPPLRSVSVKLADHATRSS
LAACLSPTTLGLVSTLVRGEGGERGAAGEAAGLPV"
                     join(5393..5402,5526..5634,6849..6936,7719..8195)
                     /gene="P0703B11.2"
                     join(5393..5402,5526..5634,6849..6936,7719..8195)
                     /gene="P0703B11.2"
                     /note="hypothetical protein"
                     /codon_start=1
                     /protein_id="BAB85284.1"
                     /db_xref="GI:18844814"
                     /translation="MYATSRKDFVTLRAQWLLAASNRNKPSTYSFLGCTTCPLV
GVQGWEDLCRRRRPSELGLEIESRCRFFOSLLLVLANPAAVAFTEDAKCPWPGGR
QRAHAKSCSPYLTQPPPTVPVVRREACGGHGGGSGRRSVRCPPVPVPPDDVA
EEEEKSLAAEEEEKLTEDGBEEQITERBELVAREEDTQTDEDDTQKTRPARPKPNP
KVHGAKWVA"
                     8569..13160
     gene              /gene="P0703B11.3"
     misc_feature       8569..13160
                     /gene="P0703B11.3"
                     /note="probably inactive due to frameshift(s) in CDS
probably inactive due to stop codon(s) in CDS
pseudogene, similar to polyprotein"
                     /pseudo
                     join(13863..13934,15133..15480)
     gene              /gene="P0703B11.4"
     CDS                join(13863..13934,15133..15480)
                     /gene="P0703B11.4"
                     /note="hypothetical protein"
                     /codon_start=1
```

```
/protein_id="BAB85285.1"
/db_xref="GI:18844815"
/translation="MEPRKGAAGLLARPRVAHLGRVYTIQIWIISDVAMHTSGWAKI
KRKESPRLOQSNPSHRYACKTARRGADGVGRKAMARFLREBNPCGAGDEGDRS
GAARHPQPPQEGGGELVAGSYSSVPVFLSLAQH" .22934)
/gene="P0703B11.5"
join(18955..19061,19338..19445,22352..22934)
CDS
/gene="P0703B11.5"
/translation="19061,19338..19445,22352..22934)
/feature="hypothenical protein"
/codon_start=1
/protein_id="BAB85286.1"
/db_xref="GI:18844816"
/translation="MVKRSDNMDSSSECSRGAKRRLQDSRSYQENAMKKVICGTRT
EYVAPYHDGOWKYGOKMTRGNSFRPCYVRCYTHQDHGCPASKVHQNSDEPPLF
RVIYTNBHTCCTNSADSYMASMOIQIADIADSLRKAQAEERLKAETVRLMHSPP
PRCSGVNMMANKERKDVIVSLLTVIRGCHTAEAGNNSAAALPNRPPPAVARSDHY
SCSAISPELLUPASDDTLDFMLDSVDFPHWVEFLDLAWLKSTHTG"
join(23961..24007,24256..24442,25834..26617,27725..27794,
29077..29191)
/gene="P0703B11.6"
join(23961..24007,24256..24442,25834..26617,27725..27794,
29077..29191)
/feature="contains EST D15357(C0514)
unknown protein"
/codon_start=1
/protein_id="BAB85287.1"
/db_xref="GI:18844817"
/translation="MVSQFPDSDNNDIRKICRIRVRSYSLIRIRDSPLYFPYS
SPYFPTPKMNTKMRALRSIRSFIRKEGLHCSATIKTSAGVWHCTGAPAG
SPHATPGSASTITRCAAITHRPSSTPAKTRPQOLPAATPSARHRLRRGLP
AEPTSPANRHOLLKTRPSAHVATSAASAACPSATVTVSGASHPVSPASAP
HRRQPPSPSIHLLQSSGQIOIPOSQRTSPATPGCATVRROOKTPPPATYLPDPC
ADLVYTSVTITSRKCHLQTSPPHRTTPMPTSSPLDPAGRHGSRPRRERVA
PLPRKLSSNDFLAGQLSWLTANVHKRDESVIDLAGHESEEDYFVVYSELO
TAISTTVR"
join(29568..29636,30926..31120,31516..31684,32551..32689,
33006..33072)
/gene="P0703B11.7"
join(29568..29636,30926..31120,31516..31684,32551..32689,
33006..33072)
/feature="hypothenical protein"
/codon_start=1
/protein_id="BAB85288.1"
/db_xref="GI:18844818"
/translation="MQLCORDLPQKSPYACDASAVIOTATARIIRGDDGGMHAPRE
CGGTLTKRENSGDEDSWEARAKVDETVFGGSLTTTGVKTSWTFGVGDGN
VLDTATELEAFRRTIPKTISSDFCLSCSGASDCHMPLADREGADACYLAAAV
TETSKLGDDDDVQAALLQIRCFPPATGARGAAVFPVVGKSTFVGFPTGY"
join(34279..34442,39279..39383,40143..40713)
/gene="P0703B11.8"
join(34279..34442,39279..39383,40143..40713)
/feature="hypothenical protein"
/codon_start=1
/protein_id="BAB85289.1"
/db_xref="GI:18844819"
/translation="WASPLKREOSFQPEEASQAEVGSASYSPPGGGVFGISPP
ESSPRDRKRRKDRPSLVKHTTFPHFDGLWRKYQKNIKDSAFPRCSYRDRCLAS
KLVQOENDDPPLRYQVYTHNTTPTVPDVAEOPPPGAGDYLALRFGSSAGG
GGGAHQOQTRERQQNTARRPFMLSFDSVSHQLEHQHPAFPDQQLPATFAAA
SPSFSFTAELAAPLTTTMDNGDGLFSTWDALRYGLDYDGHGLGNHYLPDDCNGD
DNY"
complement(join(41449..41706,42198..42595,42641..42848))
/gene="P0703B11.9"
complement(join(41449..41706,42198..42595,42641..42848))
/feature="contains EST AU094778(E31622)
unknown protein"
/codon_start=1
/protein_id="BAB85290.1"
```

```
/db_xref="GI:18844820"
/translation="MDGDGYGDSGTRDFMSQSDSHADHDELDPYTDLSYTYGRSGS
ASRIGMAALDINSGGDWPLDGSYEGGHLWGAGHGRGVGSGSHRSASQGGGADLP
PVRASVSTGLRSARPAPCRGATAPRANSPYGGRSAPAAAGRCGGRTQATGA
RCGGRSTSTAAVSEVNLDDSDNVEDAAELOEIQNHGKVPKPDRAWNTEDTSL
CEIWDQINKGNCVNGVMQKTGMREIATAHAATNLHRGSIQINRIRQLKGLWQFIQ
TLRNTQFNRI"
join(43679..43849,43967..44029,45309..45425,45617..45793)
/gene="P0703B11.10"
join(43679..43849,43967..44029,45309..45425,45617..45793)
/feature="hypothenical protein"
/codon_start=1
/protein_id="BAB85291.1"
/db_xref="GI:18844821"
/translation="MKEYGVISHOVWOLLWLLAPPGVOLFGGGMKALTRDGLWRL
TMAGPYOHFGRRLSGSLSWFRGARSYLAAGVAVNSRAOITILLDHOITIDROVGEKL
TPSHHLLTFCMFRIQIRHIIKAGLLWTYSLFYFYAFADIAMLLSSFYGTGSLSD
VSSGSEYKSKSTDTSV"
join(46353..46642,46739..46876,47886..48585)
/gene="P0703B11.11"
join(46353..46642,46739..46876,47886..48585)
/feature="hypothenical protein"
/codon_start=1
/protein_id="BAB85292.1"
/db_xref="GI:18844822"
/translation="MAFGQDTIEOLYRELAGRRLSAKLOALEPGLDSRGOKEAVDV
SRELGRVMVSLYMLKPCSSNRSEPGVTRTAPETRTDDSLCLHTPARVKVRSEVL
VRNGREVVTRTEIITSPYKDGQWRYGQKNIQDSNYLRYFKCTFSRERSCAKX
QVQDAGEPPMFLVTVNHEHTCOQPAQVGTPTNAGSSPTTTRQROSSPPPAEML
DLTMNAGLSRLLLPHAVGGSGAAEEAAIVTCLAAVISGGGAAAAPPLIPWPSA
PEAAVASAGHSPSADESVADEAAAQWADMVDFCQYQDSTFGAAAAADHRVLIG
DGDGVQRIVAARIADTVPRYTRDTSWETAGTSSMRGSID"
Query Match 7.5%; Score 107.6; DB 8; Length 144074;
Best Local Similarity 62.0%; Pred. No. 6.6e-10;
Matches 170; Conservative 0; Mismatches 104; Indels 0; Gaps 0;
QY 853 ACAGCCTTAAGGCTTCTTGATCCTCTAGTTGAGGTTGATTTTCAAACGCATGATAAAGC 912
DB 4806 ATAACTTAAGCCCTTCGATCTGAGTTGAGACAGATTTTCAGCGCACGCAAAACG 4865
QY 913 AGAAAGCTATTAGCACATTAATTACTTAGATATTATTAATAAACTTGAATAAATAT 972
DB 4866 AGAAAGCTATTAGCATATGATTAATTAAGTATTATTAATAAAATTTGAAAAATGGGT 4925
QY 973 TTATTTGAATTTTAAACAATGATGCATAAATTTTAAAAACACACCAATTTAA 1032
DB 4926 TTATTTGAATTTCTTAGACAATCTTCTAGTAAACAATTTTGTCAAAGAAGCACTTTAA 4985
QY 1033 CCTTTAAAAAGCATCTTAATAGGAAACGAGGAAGTTAAAGATTCCACCGAAGTGTTTGA 1092
DB 4986 TAATTTAAAAAGCATGTTTAAACGAAACGAGGAGTTGAGAAAAAGCACTGGGCCTTAGTC 5045
QY 1093 TAATGAAAAATGGGTGGGATTAGAAATTCGTTAA 1126
DB 5046 AAACCTTGCTAAGTTTGACTTTAAAAAATAGTCAA 5079
```

```
RESULT 12
AB023482
LOCUS
DEFINITION
AB023482
VERSION
AB023482.2 GI:5091496
KEYWORDS
SOURCE
Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
clone:P0680A03.
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
```

AB023482 156054 bp DNA linear PLN 25-MAY-2002  
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 6,  
clone P0680A03.

REFERENCE AUTHORS TITLE	Erhrhartoidaeae; Oryzaeae; Oryza. 1 Sasaki, T., Nagamura, Y. and Yamamoto, K. Oryza sativa nipponbare (GAS) genomic DNA, chromosome 6, PAC clone: P0680A03
JOURNAL REFERENCE AUTHORS TITLE	Published Only in Database (1999) 2 (bases 1 to 156054) Sasaki, T., Nagamura, Y. and Yamamoto, K. Direct Submission
JOURNAL	Submitted (05-FEB-1999) Takuji Sasaki, National Institute of Agrobiological Resources, Rice Genome Research Program; Kamondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan (E-mail: tsasaki@abr.affrc.go.jp, Tel: +81-298-38-7441, Fax: +81-298-38-7468)
COMMENT	On Jun 17, 1999 this sequence version replaced gi:4521152. Detailed information of this PAC DNA sequence is available at http://www.dna.affrc.go.jp:82/ The orientation of the sequence is from SP6 to T7 of the PAC clone. Genes were predicted from the integrated results of the following: GENSCAN1.0, BLASTN1.4, BLASTX2.0 as well as SplicePredictor (October 1998 version). The genomic sequence was searched against the non-redundant database NRP (PIR, SWISSPROT, GENPEPT, PDB) from MAFF DNABank and the cDNA sequence database at RGP. Protein similarities of the coding regions were searched against NRP with BLAST2.0. ESTs represent the identified cDNA sequences using BLASTN1.4 with the corresponding DDBJ accession no. and RGP clone ID. Sequence updated (08-Mar-1999) Sequence updated (10-Jun-1999). Location/Qualifiers 1. .156054 /organism="Oryza sativa (japonica cultivar-group)" /cultivar="Nipponbare" /db_xref="taxon:39947" /chromosome="6" /clone="P0680A03" Join(1215).1653,1694..1973,2316..2405,2407..2869) /note="Similar to Arabidopsis thaliana mRNA for DREB1B, complete cds.(AB007788)" /codon_start=1 /protein_id="BAA78732.1" /db_xref="GI:5091497" /translation="MSGGAATVAPGWRRELALGILLVPLVSAMEARGCGHHRRDLDA SAIKLSVMTVRSQGRQRRYARMSPSPPTLNCRLRRTRHGEHRQRAI DILLGGDAVVAEAGAGNOVGEAPSPVGMAMVAGVGVOGTDTAEETARTHD RAMALCAGASALNFADSAWLLHVPRAVPSVLPPAARCAATCLOGHRRVPAFGRGS TATATATGDAASTAPPAPVLSSHASSSMLATSVQNLRLATSSHLSPPSHERTMP SGAESLAPLLPFRSARLALLSPPLHRRARPPTPPRAVLPFVSHVSAASALLPHRQA VKPPCSHVAAAMQVETGAPGGGAQRSPPIGWLSHSLPAAALLPQPSVHRHSP AGCAPREKSRGKERRRERGWERGSADVAS" Join(4728..4859,4958..5037,5136..5280,5374..5799) /note="ESTs AU058067(E20733), AAU058070(E20873) correspond to a region of the predicted gene Similar to Populus tremuloides caffeoyl-CoA 3-O-methyltransferase mRNA, complete cds.(U27116)" /codon_start=1 /protein_id="BAA78733.1" /db_xref="GI:5091498" /translation="WAEASASAAATTEQANGSGGEOKTRHSEVGHKSLKSDLLQ VLETSVTPREHECKMELEVTANHPNMLTTSDEGFNLLKLKAKKMEIGVY TQYSLALAIAPDDGTILAMDINKENYELGLPSTEXAGVAHKIDFREGALPVLDDQ VEEENHGSFDFVFDADQKNLYNHERLMKLVKGLVGYDNTLWNGSVVLPADAPM RKYRYRDFVLEMLKALAAADHRVEICQLPVGDGTTLCRRVK" Join(7067..7799,7986..8163,8282..8324) /note="Similar to Human mRNA for alpha 1 (III) collagen fragment (aa 892-1023).(X01655)" /codon_start=1 /protein_id="BAA78734.1" /db_xref="GI:5091499" /translation="WRPHISSLLPSSLCPLPILLSLSRLGKQRYVTGIRWDDRS GARAPARRGAGDCGWEERSARVRRIMRGNGHPWLLPPTGARAPILLSRSPLM SFLPGSTTLPPRGQRLRRSHPPAPLPFPHPHSGSGSRGCGGCGLLAPLD LUPPTKTTITAGSRQRRLRTSSSQPPPPPSLSASTPSNVDSIIAMGNGNAAEEDD
FEATURES source	
CDS	DGDGFEQEENETQLRYGLSPGGHCKGNKCNENIPLPQSGDAGLRCHWSLPLQNGRVAPS LPRNGRAEPLSLPCSLPRATRDVTVPGRGMVYRDRRGIFV" complement(join(8393..8800,9974..10171)) /note="Hypothetical protein" /codon_start=1 /protein_id="BAA78735.1" /db_xref="GI:5091500" /translation="MQSIIVMSCRHVTEAGLVALVGGCRRLECIINVGMVPPPSFAG LLAIRPALQIRSIPOILNANVQOGGAAGAKRNGLAARDRCSTGSGAGRWRSATVPH TGRAVPRTRITAVPHRCGLDRACATGGLDRPCGTAVDCGTGKGLGPARWVGCTV AVPEGAVARAGSAEPWRSRALVLRNRSRAHWCGLGTVATV" complement(join(11764..11781,12101..12236,12390..12536, 12622..12787,13396..13751,14014..14018)) /note="Hypothetical protein" /codon_start=1 /protein_id="BAA78736.1" /db_xref="GI:5091501" /translation="MESARGRCRSPWRGGEAGGGGGGARRGAVAAEPHLRSGPS HEEEGVGGRRGTTEELRFTTLRRPRRSPATCGEGRWSTAAETEPVOGRSGAPL LRGCVLHVLTLTTLTARGEATEPKAMAGGVIVGAPPAABEPEPHVERLPADL LAHVLSLLPSPHDLISMAGVGRWRRAVERSLAGRMRMSFAGRTGDDSTARFVRAAV NLRLDLSRSCWGCOTDOGLIRISTADCVKNLTSLSLWLAGITDNGVIOLCIRSQ" Join(14440..14578,18996..20374) /note="Hypothetical protein" /codon_start=1 /protein_id="BAA78737.1" /db_xref="GI:5091502" /translation="NDSSRDSEBRPQRKNGEETENGDNGLSKDVANELLGCV VHRKMEAEWHIPSSPPNPKSARLLPSLSITTPRPTNPTCGDLAALAMPPLSF LHLAANPTLHASPLRARRLRRLSLRTIDHSRRRAVASASAGENSTPGGPIADVE MYRGDGVMTARPPTVVVLDNDKPPGPPFPAATALRDAASLLGRVVSFAFANRH AFSHVPAWADERERRAMARAAGASPPVYSCAVGRRFPTPDLTRFRLQHQ RERNKLSRLSGKGRKFRERISGNTKYDDAARELLTPKVGYLAALRLTVVGVH VTVSDKQADHALKROVKHSCVGDWLVLDSDSDTDTTKARADLRITVVVGD GCRAIGSVADTWLPHDRVENGEVDEMLRNGTHMGFRDEEEDDDSEFVMDWDSDL DGVDDIVAIRTKLFGATMSAPDEEIMDGILGVINGDMLWSDDDEDEGYF" Join(24114..24406,27608..28262) /note="EST AU055776(G20048) corresponds to a region of the predicted gene. Similar to Arabidopsis thaliana AP2 domain containing protein RAP2.10 mRNA, partial cds.(AF003103)" /codon_start=1 /protein_id="BAA78738.1" /db_xref="GI:5091503" /translation="MDNHAVAVQQRDTAPRPARARASVSIQGGGPGWADHVALPR PRCHRRPTPCRPLPAPASSIHPHRAVLPSGAPPHVSGAGRTGGAAGLAPPHS FQEKKPKPRHAGVASPRALAPLNSARMORREATVLPPLPPQPTQPOQAAAA VRAVGGGCGGGRQYRGVWRKWKGVAEIPEPNKSRIMLGYSISTAVARAADTA VFYLGRSARLUNFDOLDGAGGAGAGAEHDRELTAAVIRKKAEEVGARVDAQHSV VGAAAPVLPQPQPPPPQRRRTKNPDLNREPTPTDSDE" complement(join(33393..33435,33541..33634,36278..36468, 36579..36833,36927..37241,38176..38432)) /note="ESTs AU033170(S4339), D41681(S4339) correspond to a region of the predicted gene. Similar to Pisum sativum mRNA for pIAA4/5.1.(X68215)" /codon_start=1 /protein_id="BAA78739.1" /db_xref="GI:5091504" /translation="MDGEAWGRMAPIRHPDRLAANLQRPGETRNTNTERGASMLT RWHDAAQIRLGDSSERKRPVHAQOMLPALIPNSHHNRCRHEHFPKSLRSRV GVDAITRAVRCAARSALYVRCATATTRNVGARGGSGKRAKACRPRCGRARQ GGHARGVPRVPTDFMLICSRAPHYITTSPLAGELERRMELBLGURLAIPSPS PATAAGSELDDLNSAFSGCKRGFEALGFKTDNDNDNGGGGDDGSDGNGNK RRKLVKPPVKCLHRRRDCGGCGGVYMEGLAIKRWKLISLGSYAEALLTLHMF PTNODGDHRRRHPYAVTVDEGDGMQVDPVPAKSVKELILV" Join(40909..40960,41227..41312,41341..41475,41867..41974) /note="Hypothetical protein" /codon_start=1 /protein_id="BAA78740.1" /db_xref="GI:5091505" /translation="MGLWKNRRRDMRDGKGASNNDOIPTGNRRDEDOELPSCTASH GNRSARGORTHTADTAPTRHDCGRGRKREPPIHRSRSGRRVDAARLDLSGLKT TFTNPARTILTHLPLPLNSNR"
CDS	

CDS	join(45269. .45340,45973. .46117,46123. .46571) /note="Hypothetical protein" /codon_start=1 /protein_id="BAA78741.1" /db_xref="GI:5091506" /translation="MRRFNQNNFEYLVLAFILIRPYTFYAEFGCKSVNLKSTSTILRD SRFAAASPOLLPPFGCCPOTKPSALRASKAATSHRSPMSPPHPSPAPPRSPF HRLLRPPRPVPIRRPROGAVLPAYVSDTAALRSPNPYDNKSGNGASACAPF LLDEIGLRQRTPASSRAPRSLTLARAHGSPCWLPRNAFRPTWRARCSGCTSP VIV" complement(join(48086. .48195,48370. .48525,48574. .48616)) /note="Hypothetical protein" /codon_start=1 /protein_id="BAA78742.1" /db_xref="GI:5091507" /translation="MQDLPOVLNVVFGCAARLVBESGATGCGCGGTEKAVGGG AARCSUAVGRRGAATQQRREKNEGVDYPAVSTYPPYDAYQILILIGSAVSG "	ORGANISM	Oryza sativa Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzoae; Oryza. 1 (bases 1 to 129838) Wing, R.A., Frisch, D., Presting, G., Wood, T., Yu, Y., Soderlund, C., Kim, H.-R., Rambo, T., Henry, D. and Simmons, J. Rice Genomic Sequence Unpublished 2 (bases 1 to 129838) Wing, R.A., Frisch, D., Presting, G., Wood, T., Yu, Y., Rambo, T., Henry, D. and Simmons, J. Direct Submission Submitted (17-AUG-2000) Clemson University Genomics Institute, Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA 3 (bases 1 to 129838) Wing, R.A., Frisch, D., Presting, G., Wood, T., Yu, Y., Soderlund, C., Kim, H.-R., Rambo, T., Henry, D., Simmons, J., Wilson, R., Minx, P., Bradshaw, H. and Du, H. Direct Submission Submitted (17-MAY-2001) Clemson University Genomics Institute, Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA 4 (bases 1 to 129838) Wing, R.A., Frisch, D., Presting, G., Wood, T., Yu, Y., Soderlund, C., Kim, H., Rambo, T., Henry, D., Simmons, J., Wilson, R., Minx, P., Bradshaw, H., Du, H., Thurmond, S.K. and Mao, L. Direct Submission Submitted (24-MAY-2001) Clemson University Genomics Institute, Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA On May 17, 2001 this sequence version replaced gi1384341. This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >=30: an attempt was made to resolve all sequencing problems, such as compressions and repeats: all regions were covered by at least one plasmid subclone or more than one M13 subclone: and the assembly was confirmed by restriction digest.
CDS	complement(join(54288. .54381,54733. .54920,55012. .55254, 55374. .55652,55743. .55940,56029. .56664,56763. .57017, 57133. .57239,57324. .61120,61231. .61720,62472. .62862, 63035. .63127,63241. .63364,64098. .64378)) /note="EST AU065533 (C2174) corresponds to a region of the predicted gene. Similar to Homo sapiens splicing factor Prp8 mRNA, complete cds. (AF092565)"	FEATURES	1. 129838 Location/Qualifiers /organism="Oryza sativa" /db_xref="taxon:4530" /chromosome="7" /clone="OSJNBb0024A20" 1713. .3591 /gene="OSJNBb0024A20.1" /note="Putative Mutator protein" /protein_id="AAK55767.1" /db_xref="GI:14192862" /translation="MDRLVRLHYGGSVVQSNPNFEGMSVKOLIIFGKPSFBEELVC HTKDLVGSNQSTAIQGRYDVAGPI SHKYMLDLNGLEWNTYVDVLGSOFSKLEVV ASKLDRDRCDDSDFLNRTLFYDEPNYETFDDEGDMKSQSVKSLTEVEVIESQMGVD VNMKKEADAKGRSQRSESEINVELQKEEAANKKSELGVEVNVVQLQKEEA IEVNEASQAVPIVGNVDODNEVADGSMGGBRETTDAQLASLRFVCFDKIYTRR ECESTLLMGLSDVLLHSEEDDTHEEGYEVEYAVDSDDPVAIKSQSEREVE KLVCNPEVQFEDVNNSELAVDGVYNGIIDAMGEDIVHIIFKLSLRHCLLKLDD RRKSKKITDMHAHLHLIERWDNTAANVPHGVHKNKTYFDEYLAWLRLNRYLFLRPA WTLADIATDDEVEEMNEYDTHTFAGATVEYGPVDRVARELMSVNDAGVALCTPAG SENEVGILRSALQKLKORSRLAARLCGRSTNAVDQPRLTRTQGETSQARABEEG" 1713. .3591 /note="Putative Mutator protein" complement(10720. .11542) /gene="OSJNBb0024A20.2" /note="Hypothetical protein" complement(join(10720. .11360,11455. .11542)) /gene="OSJNBb0024A20.2" /codon_start=1 /product="Hypothetical protein" /protein_id="AAK55768.1" /db_xref="GI:14192863"
CDS	893 TTTCAACGCATGATAACGACGAAGCTCATTAGCACATTATTACTAGATTTATTAAT 952 Db 105226 TTTTTAGCGCAGCGGAAAGAGCTCATTAACACATGATTAATAGTATTGATAAT 105285	gene	
CDS	953 TATAAACTTGAATAAATAATTTATTGATTTTAAACAAATGTATGCATAAATTTATTT 1012 Db 105286 TATAAACTTAAATAATGATTTATTGATTTTAAACAACTTCTATATAGATT-TTTT 105344	CDS	
CDS	1013 TTAATAACACACCAATTTAAACCTTTTAAACACATCTCTAATAGGAAACGAGGATTA 1072 Db 105345 TTAATAAAGATATCATTTAAACAGTTTGAACGCGTGCTAACGAAAAATGAGGAAGTTGAA 105404	repeat_region	
CDS	1073 GATT 1076 Db 105405 ATTT 105408	gene	
RESULT 13		CDS	
AC079038/c			
LOCUS	AC079038		
DEFINITION	Oryza sativa chromosome 7 clone OSJNBb0024A20, complete sequence.		
ACCESSION	AC079038		
VERSION	AC079038.3		
KEYWORDS	HTG.		
SOURCE	Oryza sativa.		



/translation="MGNWRDEGVSCSEIRVGQVSHRGASLSSEAPASEVHVLQAPDQSH  
NFCINCELPVGVGRYCNHHCSDFAIHEACADRFAFARDITDGFFAHPRIPLTAATAASAA  
RPGWCDLCOGPTQPGRLVYCAEACAFDHPVLCITLLPMTIHPAARRRVPFRPRQLGP  
GARGVGVLPGLLHIALPVRPMLVHAPHLRFPAGRRRCSSWYEGLACHRRRRLRGLR  
QWQSHREVSDHVCATGGEFDCQ"  
14044..14310  
/note="Similar to Tourist03 type MITE element"  
16254..16337  
/note="Simple Sequence Repeat (ATAC)n"  
16282..16849  
/note="Similar to Micropon3 repeat sequence"  
16338..16387  
/note="Simple Sequence Repeat (AT)n"  
16781..16852  
/note="Simple Sequence Repeat (AT)n"  
complement(24253..24966)  
/gene="OSJNBb0024A20.3"  
/note="Hypothetical protein"  
complement(24253..24966)  
/gene="OSJNBb0024A20.3"  
/codon\_start=1  
/product="Hypothetical protein"  
/protein\_id="AAK55769.1"  
/db\_xref="GI:14192864"  
/translation="WATRRRHFAHPQHLLLTHYSTTSSHFCDICRTNVAGMVGYRC  
NTCIDIDHACADYFEETVSFFAHPMHTLTSTIPDDTNRWSCDLCTESCPRGSFVY  
RCAQCQFDHPLCTLLPOTIRSLPHQHDLMNAPSWGTCSCACHQGLNMWHYRCGFLY  
KLHVPCGAANAGOGDTSRSHNAGAGAAATPGSSRATRIAKFLRLTSLNAINAAT  
GALATPVLDTALDDQE"  
27626..27972  
/note="Similar to Tourist03 type MITE element"  
complement(30242..32014)  
/gene="OSJNBb0024A20.4"  
/note="Hypothetical protein"  
complement(join(30242..30813,31891..32014))  
/gene="OSJNBb0024A20.4"  
/codon\_start=1  
/product="Hypothetical protein"  
/protein\_id="AAK55770.1"  
/db\_xref="GI:14192865"  
/translation="MCPELSILNTTSIEPRNEGGLVLKDNFVSMIPQASATENBEAAS  
GNQLESRTFHAHPQHLTKYGGGRQPSRRHVCRCGRKVAGAGRCDHCDFDI  
HEACADFPEKMITPPNFPFHPHNLALRQVTAADSGSWPCTLCRGPFGHGLAYRC  
GARRCGFAHPLCTMLPGSIRSLPHRKALHTLEMLIPSGPCTPAEMARVCSVCRD  
CSTVRTRHYRCAA"  
complement(36235..36738)  
/gene="OSJNBb0024A20.5"  
/note="Hypothetical protein"  
complement(36235..36738)  
/gene="OSJNBb0024A20.5"  
/codon\_start=1  
/product="Hypothetical protein"  
/protein\_id="AAK55771.1"  
/db\_xref="GI:14192866"  
/translation="MAASVTKLHPGHSILKRHHYGGEGHACVCAICBRIIAGAGYRC  
GSGCGFDAHEACLSLPMRVAFVGHPAHELTLSTLTSRWCDACRVASHAGRCVTRCVA  
CDFDVHARCTSLLDGEQQHGRKRGVARRVGMALRMGLFGLRVADVTVGGFGSPVIEV  
ITATNLN"  
complement(38420..42086)  
/gene="OSJNBb0024A20.6"  
/note="Unknown protein"  
complement(join(38420..38681,38768..38895,39023..39161,  
39249..39354,39784..40031,40123..40256,41108..41277,  
41891..42086))  
/gene="OSJNBb0024A20.6"  
/codon\_start=1  
/product="Unknown protein"  
/protein\_id="AAK55772.1"  
/db\_xref="GI:14192867"  
/translation="MSHYRSAPAGKSGGGGSTPAKRGPRPGSTTTGSGAAAAA  
GAAAPAAVLGSLQVLTALSDQNNKRIVLALQSLKSEILWALNALTVLSFKKDDL  
RBDTPPLAKVPGLLDALQVDDWRDIAMPKDHTRPPRVTLGYNLTLSGFHENVEK  
VYSDTTTPSDDOTKTADSTVTKRSAGFLFDEGLFNVDDEGRTEKQCAVAASNIIR

NFSFMPENETVMQVHRHCLETVFQCLDQNTDEDELITNMLETVLNVLAPVLDIRFSS  
SKSFKITKRAVQAQAIMGLASSIRVWHCAAELIGRLIINDPNDFPFLLPQIYIK  
RLVDLSVADQAQAAISALYNVAEVMNDFRKLJASERMAVDRLLKVKTPHPPEV  
CRKASIVELSVSEPQNRHMLLVHENTFABIITSEGYSDTFTARILYELTARPSNKVT  
AGQAIWGIN"  
complement(44470..45027)  
/gene="OSJNBb0024A20.7"  
/note="Hypothetical protein"  
complement(44470..45027)  
/gene="OSJNBb0024A20.7"  
/codon\_start=1  
/product="Hypothetical protein"  
/protein\_id="AAK55773.1"  
/db\_xref="GI:14192868"  
/translation="MAGVSCFYHPEHLVTEHHYGECSGSGNPCAACERVVTGDGYR  
CEHGCDPFVHRSCALPASFALHERHEHLETLTLTGSRCNVCDVISHAGCRMVAC  
APCNFSAHPRECTPVLGDAQHOQPOPOQOQEGEEKRSVGRAAKVALKIGVAVGLTAVDV  
GTTGGLATVGSLVIAPIVRADIKI"  
46231..46470  
/note="Similar to ditto\_Os1 MITE element"  
47918..54073  
/gene="OSJNBb0024A20.8"  
/note="Putative polyprotein"  
join(47918..50272,50345..54073)  
/gene="OSJNBb0024A20.8"  
/codon\_start=1  
/product="Putative polyprotein"  
/protein\_id="AAK55774.1"  
/db\_xref="GI:14192869"  
/translation="MAEQAKHDLSPSVSGDDGEPNRRARTPPPPHOSPKQGEAL  
ERVEKSAAPVAGDARERRDGERLLVYGDGSTPGQALQVALLRHPVPVPESPA  
QRWLDVYANLVMTAQORLQALGASRATKTTGSAATGTSVSRRARAAAVARSAATP  
SSAPPTREDOHEEPDARLDIERRRNSRRTPRAAEGASSRSVPRQGEDQSPVPVGG  
VGCRAFVASLRNVWP RPRTITEKYDGSVNPTFLOVYTTGIEAAGGDDRVMANFF  
PMALKQAGRWLMNLPASVSHMEDLCQFTNNFQCTYPRGEEADLHVAQCGDDEL  
RSTQRCVRNTIPCI PAHAVIYAFRGVRHNRMLEKTASKPEQITTAELFQLADRV  
RKBEAATWPNSSGSAASAAPSAATGRRRRRRKSVHSGDEGHVLAVEGAPRATR  
KGRPASDKKEAGAPRERPAQKWSVHTSLDLADCRVKNLARTKKWEEDRQE  
RRENKSPALVSGNRRREAKQPAEDIDGDDDLGFOEPCGATATVDGGAACHVSRRS  
FKAMRELLAAVPTHEATRARSEVALTFDQDTHPPCVARGGQIAMVVSPTICNVKL  
GRVLIDGGTALNLSPAAPDAIKAPGMVLBPSPQIIGVTPGHTWPLGHIDLPVTFGS  
ANFRTERKVDVADLSLPYNALVGRFALVMAAHYALQMKMPGGPGGFIHVGDUK  
VALACMEQRADHILAAASKPEGDERLGTSAPTAPRQRIPTCDQDVPDALVSFLRANAD  
VFARWADMPGVPREVIEHRLAVRPGARPVQRVQRQAPEROQAFIREEVALLEAGFI  
REVTHPELANPVVVPKANGKRCMIDYDLNACPKDPYLPRIQIDVIDSTAGCDLL  
CFLDAYSGYHQIRWAREDEKTAFTIPVGTVCYVTSMPFGLNKAGPTFOHTTRISLGSQ  
IGRNEAYVDDLKVTNRQETFLSLAEFENLRSARIKLNPKKCVFGVPAGKLLGFL  
VSARGIEANPEKIRAIERNMRPPSKLRDQVCVTCMAALSRFISRLGEKALPLPLUKR

Query Match 7.3%; Score 104.8; DB 8; Length 129838;  
Best Local Similarity 76.6%; Pred. No. 2.le-09;  
Matches 141; Conservative 0; Mismatches 42; Indels 1; Gaps 1;  
Qy 894 TTCAAAACGCATGATAAAGCAAGAGCTATTAGCACATTATTACTTAGATATTTATAATT 953  
Db 62337 TTTTAGTCACGCAAAACGAGAAAGCTCTTAGCACATGATTAATAATAATAATT 62278  
Qy 954 ATAACTTGAAAAAATAATTTTATTTGAATTTTTTAAACAATGATGCATAAATTTT 1013  
Db 62277 ATAAACTTAAAAAATAGATTTTATTTGAATTTTTTAAACAACACTTTATATAGATTTT 62218  
Qy 1014 TAAAAACACACCAATTTTAAACCTTTAAAGCATCCTCTAA-TAGGAAACGAGGAAGTTAAA 1072  
Db 62217 AAAAAACACATCATTTAAACAGTTTGAAAGCGTGCTTAACAAAAAAGGAAGATTGAA 62158  
Qy 1073 GATT 1076  
Db 62157 GTTT 62154  
RESULT 14  
OSJN00031/c  
LOCUS OSJN00031 170025 bp DNA linear HTG 12-JUL-2002





```

source
1. .131983
/organism="Oryza sativa"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:4530"
/chromosome="10"
/map="R1877"
/clone="OSUNBa0051D19"
complement(135566..15645)
/rpt_family="AT_rich"
1038..1061
/rpt_family="CAT)n"
complement(1191..93025)
/rpt_family="Oryza sativa gene, repeat sequence Micropon-3
gi|4586621|dbj|AB010113.1|AB010113"
complement(1193..1256)
/rpt_family="Oryza sativa gene, repeat sequence Micropon-2
gi|4586622|dbj|AB010114.1|AB010114"
complement(1194..93014)
/rpt_family="Oryza sativa gene, repeat sequence Micropon-4
gi|4586620|dbj|AB010112.1|AB010112"
complement(1195..1254)
/rpt_family="(TA)n"
2878..3430
/gene="OSUNBa0051D19.1"
/note="EST AU089716 from this gene"
2878..3430
/gene="OSUNBa0051D19.1"
3115..3430
/gene="OSUNBa0051D19.1"
/codon_start=1
/product="unknown protein"
/protein_id="AAG13537.1"
/db_xref="GI:10140703"
/translation="WTLRLSLPDLAWPRDPAAGVVAQRRRRYGVORRGAPLWR
RRLLRPARAPTHASGARGVARSRRRGISAAHLRLGRHCPSPRRGVRPLRPAPP
SP"
5340..7117
/gene="OSUNBa0051D19.2"
/note="predicted by genemark hmm"
join(<5340..5405,6473..7117)
/gene="OSUNBa0051D19.2"
join(5340..5405,6473..7117)
/gene="OSUNBa0051D19.2"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAG13553.1"
/db_xref="GI:10140719"
/translation="MANGGELRLAASPGLRWKRTDFVSVVAFIITLTPAPLISFPNS
LSGDKVRAPIGEURTRAGMARRAASRALRSEGIQGRGAGGGGGAEDARHVFDE
LLRRGASIVGLNRLADVARGSPAAVSYNRMARAGADEVTPDICTYIGLIGCC
RAGRLDLGFAALGNVKKGRFVEAITFTLLKGLCADKRTSDAMDIVLRMRTELGCIP
PEMKQQLTFLHKHFQCLL"
complement(10362..10411)
/rpt_family="AT_rich"
12170..12246
/rpt_family="(GA)n"
12873..13312
/gene="OSUNBa0051D19.3"
/note="predicted by genemark hmm"
join(<12873..13087,13192..13312)
/gene="OSUNBa0051D19.3"
join(12873..13087,13192..13312)
/gene="OSUNBa0051D19.3"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAG13530.1"
/db_xref="GI:10140696"
/translation="MRWLTSSACCGRHLLHAASGPHVLVIVETRIHAQHEKPPRHQSC
LSRRRLHAAHVVRTVSGSRLRPHKRGASVNTSWPASLRWSVTSVTSFGLTEDMDD
MNTSLMAVL"
complement(13487..13511)
/rpt_family="AT_rich"

repeat_region
complement(13523..13556)
/rpt_family="AT_rich"
repeat_region
complement(13576..13659)
/rpt_family="(A)n"
repeat_region
14146..15227
/rpt_family="retroelement from gi|5803242|AP000399 Oryza
sativa (75139 to 88038) 12900 nt"
complement(15566..15645)
/rpt_family="AT_rich"
complement(16948..16968)
/rpt_family="(CGAG)n"
complement(16971..20492)
/gene="OSUNBa0051D19.18"
/note="highly similar to alpha galactosidase GI:927577
(Phaseolus vulgaris)"
complement(join(<16971..17281,17368..17458,17805..17887,
17999..18063,18166..18249,18430..18495,18601..18659,
18815..18890,19085..19188,19323..19357,19691..19747,
20076..20137,20344..20492))
/gene="OSUNBa0051D19.18"
complement(join(16971..17281,17368..17458,17805..17887,
17999..18063,18166..18249,18430..18495,18601..18659,
18815..18890,19085..19188,19323..19357,19691..19747,
20076..20137,20344..20492))
/gene="OSUNBa0051D19.18"
/codon_start=1
/product="putative alpha-galactosidase"
/protein_id="AAG13536.1"
/db_xref="GI:10140702"
/translation="MAAAAVLLAVACAAAMTAGCGGRVHVHGEHRRSMLANGIA
TAPPMGWNHFPACDNGEDVIRETADALVSTGLAAAGKYVNIIDCWAEPODRTSG
NLVANKKTPFGIKALADYVHSLKGIYSDAGFKTCAKVPGSLGHEQDARTFAS
WGVYLYKYNCCNNGDLPLERYPEMSKALMAGRPYFSLCEWGDHMKPAGAAVNS
WRTNIDIAITWESVNRADENYVADYAPGWNDDPMLFVNGVNDDEYIVHFSIW
AISKAPLIGCDVRHMSOETYDILANKVITAVNQDALGIQKKVMEGSESVAAPLS
GYTAVLLNHAAREAAITAHWDDVGLPAGTAVEARDLWRHATVAGAGSELTEKITL
DVAPHSCRMILLKPLISSRN"
complement(19366..19387)
/rpt_family="AT_rich"
complement(20426..20496)
/rpt_family="(CGG)n"
22074..25140
/gene="OSUNBa0051D19.4"
/note="EST AU056178, AU056179 from this gene"
join(<22074..22556,23205..23324,23408..23461,23824..23987,
24380..24556,24638..25140)
/gene="OSUNBa0051D19.4"
join(22074..22556,23205..23324,23408..23461,23824..23987,
24380..24556,24638..24848)
/gene="OSUNBa0051D19.4"
/codon_start=1
/product="unknown protein"
/protein_id="AAG13540.1"
/db_xref="GI:10140706"
/translation="WTFPHLPVSSSSSHLLDAAPTASPAAHQHRRRRRRRCGGGRG
RGGAHQVRCRAAAAPQPVATGAGLRGAATRVFVSDLHTDIRENMDVLRPLVG
GGGGGGGVGIDALVAGDVETRDNFARTMAALRRERFVFPVGNHDLWLRNGR
YMDLSEKLTALLDACSELGVDTGPRIIGDLIGIIPFVSWYHKSFDEKDVNSVRVPSLE

Query Match 7.1% Score 102; DB 8; Length 131983;
Best Local Similarity 68.6%; Pred. No. 6.6e-09;
Matches 152; Conservative 0; Mismatches 70; Indels 1; Gaps 1;

Qy 851 ACACAGCCTTAAGGCTTCTTGTATCTCTAGTTGGAGGTGATTTTCAACCCCATGATAAA 910
|||||
Db 78387 ACACATCAAGGACCTGTTCGATCTCTCAGTTGAGAGAGAAATTTCCAGCGCACATATA 78446
|||||

Qy 911 CGAGAAAGCTCATTAGACACATTTACTTAGATATTTTATAATTAACATTTGAAAAAAT 970
|||||
Db 78447 CGAGAAAGCTCATTAGACACATGATTAATTAATTAATTTGTTACAAAATTTG 78506
|||||

Qy 971 ATTTATTGAAATTTTTTAAACAATGTATGCATAAATTTATTTTAAAAAACACACCAATTT 1030
|||||

```

Db 78507 ATTCATTGAA-TTTTAAACAACCTTTACATAACTTTTTTTTATAAAAAACACACTATTT 78565

Qy 1031 AACCCTTTAAAAAGCATCCTAATAGGAAACGAGGAAAGTTAAAGATT 1076

Db 78566 AACCGTTTGAAAAAACGATGATAATGGAAACCGAGGAAAGTTGTAATTT 78611

Search completed: March 15, 2003, 15:53:37  
Job time : 5046.28 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

```
Run on:      March 15, 2003, 09:09:00 ; Search time 268.521 Seconds
            (without alignments)
            12076.811 Million cell updates/sec
```

**Title:** US-09-802-937-2

Perfect score:

Sequence: 1 ctgcaggaagattaattag.....cgcgacgcgaattcgtcgcac 1440

Scoring table: IDENTITY NUC

scoring card: IDENTIFY\_NOC  
Gapop 10.0 : Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

1:	/SID27/cgdata/geneset/geneseqn-emb1/NA1980.DAT.*
2:	/SID27/cgdata/geneset/geneseqn-emb1/NA1981.DAT.*
3:	/SID27/cgdata/geneset/geneseqn-emb1/NA1982.DAT.*
4:	/SID27/cgdata/geneset/geneseqn-emb1/NA1983.DAT.*
5:	/SID27/cgdata/geneset/geneseqn-emb1/NA1984.DAT.*
6:	/SID27/cgdata/geneset/geneseqn-emb1/NA1985.DAT.*
7:	/SID27/cgdata/geneset/geneseqn-emb1/NA1986.DAT.*
8:	/SID27/cgdata/geneset/geneseqn-emb1/NA1987.DAT.*
9:	/SID27/cgdata/geneset/geneseqn-emb1/NA1988.DAT.*
10:	/SID27/cgdata/geneset/geneseqn-emb1/NA1989.DAT.*
11:	/SID27/cgdata/geneset/geneseqn-emb1/NA1990.DAT.*
12:	/SID27/cgdata/geneset/geneseqn-emb1/NA1991.DAT.*
13:	/SID27/cgdata/geneset/geneseqn-emb1/NA1992.DAT.*
14:	/SID27/cgdata/geneset/geneseqn-emb1/NA1993.DAT.*
15:	/SID27/cgdata/geneset/geneseqn-emb1/NA1994.DAT.*
16:	/SID27/cgdata/geneset/geneseqn-emb1/NA1995.DAT.*
17:	/SID27/cgdata/geneset/geneseqn-emb1/NA1996.DAT.*
18:	/SID27/cgdata/geneset/geneseqn-emb1/NA1997.DAT.*
19:	/SID27/cgdata/geneset/geneseqn-emb1/NA1998.DAT.*
20:	/SID27/cgdata/geneset/geneseqn-emb1/NA1999.DAT.*
21:	/SID27/cgdata/geneset/geneseqn-emb1/NA2000.DAT.*
22:	/SID27/cgdata/geneset/geneseqn-emb1/NA2001A.DAT.*
23:	/SID27/cgdata/geneset/geneseqn-emb1/NA2001B.DAT.*
24:	/SID27/cgdata/geneset/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
5	100.0	100.0	100.0	422	24	ABL09560	Rice acetylcholinesterase
5	100	100	100	422	24	ABL49953	Rice Wanderer-Osl
6	97.4	6.8	423	24	ABL49955	Rice Wanderer-Osl	
7	97.4	6.8	424	24	ABL49954	Rice Wanderer-Osl	
C 8	66.6	4.6	5875	24	ABL32289	Human immune system	
C 9	64.8	4.5	6048	24	ABQ67002	Human angiogenesis	

## ALIGNMENTS

## RESULT 1

RESOL 1  
AAA07482

ID AAA07482 standard: DNA: 1440 BP.

XX

AC AAA07482;

XX

DT 17-JUL-2000 (first entry)

XX

DE Rice adenylate kinase DNA sequence fragment #2.

XX Rice; adenylate kinase; promoter; structural gene expression;  
KW plant gene expression; herbicide resistance; stress resistance;  
KW light resistance; antibacterial production; flowering induction;  
KW insecticidal substance production; ss.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

OS *Oryza sativa*.



PN WO200015812-

XX

PD 23-MAR-2000.

XX

PF 08-SEP-1999; 99WO-JP04847.

XX

PR 10-SEP-1998; 98WO-JP04088.

[illegible]

PA (NISC) NISSAN CHEM IND LTD.  
yy

XX  
DT  
Uchimizu W  
Kaji C  
Eckhart E  
Eckert M  
Eckmann H

РІ ОСНОВНІ НАДАНІ, НАДАНІ  
УУ

WBT: 2000-271449/23

DK 2000-2/1440/23.  
XX

PT function in plants for controlling expression of desired foreign  
PT structural genes, useful for herbicide resistance and in sucrose  
PT induction in recombinant plants -

XX Claim 10; Page 28; 33pp; Japanese.

CC This sequence represents a fragment of the rice adenylate kinase gene.  
CC The invention relates to a DNA fragment (such as this sequence), that has  
CC a regulatory function (preferably acting as a promoter) on the expression  
CC of desired foreign structural genes in a plant when introduced into a  
CC plant body or plant cell. The DNA fragment can be inserted into a plant  
CC to control the expression of desired foreign structural genes, useful for  
CC inducing in the plant e.g. herbicide, stress and light resistance,  
CC production of antibacterial and insecticidal substances, sucrose  
CC production in recombinant plants, and in promoting flowering and crop  
CC yield. The DNA fragment has a regulatory function on the expression of  
CC desired structural genes in a plant, particularly with control of the  
CC expression in specific tissues and locations to provide safe chemicals.

XX Sequence 1440 BP; 506 A; 246 C; 272 G; 416 T; 0 other;

Query Match 100.0%; Score 1440; DB 21; Length 1440;  
Best Local Similarity 100.0%; Pred. No. 9.6e-288;  
Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAGGAGATTAAATTAGGTGGACACACCAAAACCCCTGTGGTGGTGACGCCCTGTGTG 60  
DB 1 CTGCAGGAGATTAAATTAGGTGGACACACCAAAACCCCTGTGGTGGTGACGCCCTGTGTG 60  
QY 61 AATCAACTGGGGTTCGTTGGACATGGTTTTTGCAGGAAATTAAGCAAGAAATTAAG 120  
DB 61 AATCAACTGGGGTTCGTTGGACATGGTTTTTGCAGGAAATTAAGCAAGAAATTAAG 120  
QY 121 AAGAATGCTCAAGCTGACATGAGAAACGTAATCCCAATGAAGCGGAATTTCAAGTCGTTTC 180  
DB 121 AAGAATGCTCAAGCTGACATGAGAAACGTAATCCCAATGAAGCGGAATTTCAAGTCGTTTC 180  
QY 181 TCTTGACTACCTACCTGTTGATACATAGACAGTGCACACGTTTGAATGCTCCTATTGG 240  
DB 181 TCTTGACTACCTACCTGTTGATACATAGACAGTGCACACGTTTGAATGCTCCTATTGG 240  
QY 241 CTGCTGTGATACCTGCTGTCACAAAGCATCAATGCTTCTTGAGTATCTTTATTA 300  
DB 241 CTGCTGTGATACCTGCTGTCACAAAGCATCAATGCTTCTTGAGTATCTTTATTA 300  
QY 301 CCGAAACCCCAAGATTATTTCTTCAAAAAGAAATTAATTAATTTTTCATGAAA 480  
DB 301 CCGAAACCCCAAGATTATTTCTTCAAAAAGAAATTAATTAATTTTTCATGAAA 480  
QY 481 CGCAATTCACCGTTCGAGAAATATGCTGTCATAAATAAGTAGTGTAGTGCAGAAACAA 540  
DB 481 CGCAATTCACCGTTCGAGAAATATGCTGTCATAAATAAGTAGTGTAGTGCAGAAACAA 540  
QY 541 ATTAATATCATATAAAGAGGTTGTTAATTACAAACCATGTTTCGTACTACACTCT 600  
DB 541 ATTAATATCATATAAAGAGGTTGTTAATTACAAACCATGTTTCGTACTACACTCT 600  
QY 601 AATTTGTAAATCTTATTTTCAGTCACAAAATTCCTCAATTTTCAAGTAAAGAAATAAACGT 660  
DB 601 AATTTGTAAATCTTATTTTCAGTCACAAAATTCCTCAATTTTCAAGTAAAGAAATAAACGT 660  
QY 661 AGACGGCTAAGCCCACTCTAAGGCTAAGTTCGAGAGTGAAGTACGACGCAAAAATA 720  
DB 661 AGACGGCTAAGCCCACTCTAAGGCTAAGTTCGAGAGTGAAGTACGACGCAAAAATA 720

QY 721 TGATGGTTTAAATATATGATTTTTTTTAAATAAATTTTACATAAATTTCTTTAGGAAACA 780  
DB 721 TGATGGTTTAAATATATGATTTTTTTTAAATAAATTTTACATAAATTTCTTTAGGAAACA 780  
QY 781 TATCATTTAAATGGTTTGAATAAATCGTGCAATAGAAACATAGAAACGATGAGTTGGGAAA 840  
DB 781 TATCATTTAAATGGTTTGAATAAATCGTGCAATAGAAACATAGAAACGATGAGTTGGGAAA 840  
QY 841 CAAGAGAAACACACAGCCTTAAGGCTTCTTATGATCTCTAGTTGGAGGTTGATTTTCAAC 900  
DB 841 CAAGAGAAACACACAGCCTTAAGGCTTCTTATGATCTCTAGTTGGAGGTTGATTTTCAAC 900  
QY 901 GCATGATAAAGCAGAAAGCTCATTTAGCACATTTACTTAGATATTTATAATATAAAT 960  
DB 901 GCATGATAAAGCAGAAAGCTCATTTAGCACATTTACTTAGATATTTATAATATAAAT 960  
QY 961 TGAATAAATATTTTAAATTTTAAACAATGATGATGATGATGATGATGATGATGATGAT 1020  
DB 961 TGAATAAATATTTTAAATTTTAAACAATGATGATGATGATGATGATGATGATGATGAT 1020  
QY 1021 ACACCAATTTAAACCTTTTAAAGAGCATCTTAATAGGAAACGAGGAAATTAAGATTCACC 1080  
DB 1021 ACACCAATTTAAACCTTTTAAAGAGCATCTTAATAGGAAACGAGGAAATTAAGATTCACC 1080  
QY 1081 GAAGTGTGGATAATGAAATGGGGTGGGATGAGATTTGATGATGATGATGATGATGATGAT 1140  
DB 1081 GAAGTGTGGATAATGAAATGGGGTGGGATGAGATTTGATGATGATGATGATGATGATGAT 1140  
QY 1141 GATTAAATATTTAAATGAAAGGAGGAGATGAAATGGTTAGAGTTTAAATGTTCTTTTGG 1200  
DB 1141 GATTAAATATTTAAATGAAAGGAGGAGATGAAATGGTTAGAGTTTAAATGTTCTTTTGG 1200  
QY 1201 GTGGGTAGAAAATTTATTTGCCATACATCTCCCGAGAGCGGTGCGTGTGCGTGGGCGAGA 1260  
DB 1201 GTGGGTAGAAAATTTATTTGCCATACATCTCCCGAGAGCGGTGCGTGTGCGTGGGCGAGA 1260  
QY 1261 AGCGTCTTTTGGTGGAAAAAATCTGCTTAAAGAGAAACAGAAAGAGCCGAGCTTT 1320  
DB 1261 AGCGTCTTTTGGTGGAAAAAATCTGCTTAAAGAGAAACAGAAAGAGCCGAGCTTT 1320  
QY 1321 GTTGTACCGTCTCACAGAAACGAAACGAAACGAAACGAAACGAAACGAAACGAAACGAAAC 1380  
DB 1321 GTTGTACCGTCTCACAGAAACGAAACGAAACGAAACGAAACGAAACGAAACGAAACGAAAC 1380  
QY 1381 GACCGAGACTCTCTCATTTTACGCGCGCAGCGGAGAGCAGCGGAGAGTTCGTCGAC 1440  
DB 1381 GACCGAGACTCTCTCATTTTACGCGCGCAGCGGAGAGCAGCGGAGAGTTCGTCGAC 1440

RESULT 2

AAAA10961  
ID AAA10961 standard; DNA; 1440 BP.

XX AAA10961;

AC AC  
XX AC  
DT 14-JUL-2000 (first entry)

XX Partial rice adenylate kinase gene promoter fragment.

DE Rice; adenylate kinase; promoter; gene expression control; antibacterial;  
KW insecticide; light resistance; sucrose induction; flower promotion;  
KW increase crop yield; herbicide resistance; ds.

XX Oryza sativa.

XX WO200015811-A1.

XX 23-MAR-2000.

PF 10-SEP-1998; 98WO-JP04088.

XX 10-SEP-1998; 98WO-JP04088.

XX



(NISC) NISSAN CHEM IND LTD.

Uchimiya H, Arai S, Fushimi T, Tagawa M;

WPI; 2000-271447/23.

Rice adenylate kinase gene-originated DNA fragment with promoter  
function in plants for controlling expression of desired foreign  
structural genes, useful for herbicide resistance and in sucrose  
induction in recombinant plants -

Example 4; Page 26; 30pp; Japanese.

This sequence represents a partial rice adenylate kinase gene-originated  
DNA fragment with promoter function in plants. The fragment can be used  
for controlling the expression of desired foreign structural genes. The  
DNA fragment is used to create a vector, which in turn is used to  
transform bacteria or plant cells. The DNA fragment can be inserted into  
a plant to control the expression of desired foreign structural genes,  
useful for inducing in the plant e.g. herbicide, stress and light  
resistance, production of antibacterial and insecticidal substances,  
sucrose induction in recombinant plants, and in promoting flowering and  
crop yield.

Sequence 1440 BP; 506 A; 246 C; 272 G; 416 T; 0 other;

Query Match 100.0%; Score 1440; DB 21; Length 1440;  
Best Local Similarity 100.0%; Pred. No. 9,6e-288;  
Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCGAGGAATTAATAGGTGGACACACCAACCCCTGTGTTGGTGACGCCCTGTTGTT 60  
DB 1 CTGCGAGGAATTAATAGGTGGACACACCAACCCCTGTGTTGGTGACGCCCTGTTGTT 60  
QY 61 AATCACTGGGTGTTGTTGACATGTTTTTGCAGGAAAATTAAGCAAGAAAATTAAG 120  
DB 61 AATCACTGGGTGTTGTTGACATGTTTTTGCAGGAAAATTAAGCAAGAAAATTAAG 120  
QY 121 AAGAATGCTCAAGCTGACATGAGAAAACGTAATCAATGGAAGCGAATTTCAAGTCGTTTC 180  
DB 121 AAGAATGCTCAAGCTGACATGAGAAAACGTAATCAATGGAAGCGAATTTCAAGTCGTTTC 180  
QY 181 TCTTGTAACCATGTTTGAATACATGAACAGTGCACAGCTTTGATGGCTCTATGG 240  
DB 181 TCTTGTAACCATGTTTGAATACATGAACAGTGCACAGCTTTGATGGCTCTATGG 240  
QY 241 CTCGTGTGATGACTGTTGTGCAAAAGCATCAATGCTTCTTGAGTATCTTTATTA 300  
DB 241 CTCGTGTGATGACTGTTGTGCAAAAGCATCAATGCTTCTTGAGTATCTTTATTA 300  
QY 301 CCGAAAACCCCAAGATTAATCTATTCACCTCAGGGTAATTTGTGCTGAATGCAATG 360  
DB 301 CCGAAAACCCCAAGATTAATCTATTCACCTCAGGGTAATTTGTGCTGAATGCAATG 360  
QY 361 AATACAAATTCGCAAAATATCATGTTATCTATCTGTCAAAATGAAATTTGAGTCCAA 420  
DB 361 AATACAAATTCGCAAAATATCATGTTATCTATCTGTCAAAATGAAATTTGAGTCCAA 420  
QY 421 CTGAGACTGCAATACGATTTTCTTTTCAAAAGAAAATTAATTAATTTTTCATGAAA 480  
DB 421 CTGAGACTGCAATACGATTTTCTTTTCAAAAGAAAATTAATTAATTTTTCATGAAA 480  
QY 481 CGCAATTCACCGTTCGAGAAATGCTGTCATTAATAGTAGTCTAGTCAGAGAAACAAA 540  
DB 481 CGCAATTCACCGTTCGAGAAATGCTGTCATTAATAGTAGTCTAGTCAGAGAAACAAA 540  
QY 541 ATTAATATCACATAAAAGAGGTTGTTTAATTAACAAACCACTGTTTCGTACTACAACCTCT 600  
DB 541 ATTAATATCACATAAAAGAGGTTGTTTAATTAACAAACCACTGTTTCGTACTACAACCTCT 600  
QY 601 AATTGTAATTTCTTATTTTCAGTCACAAAATTCGAATTTCCAAATTAAGAAAATAAAGCT 660  
DB 601 AATTGTAATTTCTTATTTTCAGTCACAAAATTCGAATTTCCAAATTAAGAAAATAAAGCT 660

QY 661 AGACGGCTAAGCCCAACCCATCTAAGCTAAGTTCGAGAGGTGAAGTACGACGAAATA 720  
DB 661 AGACGGCTAAGCCCAACCCATCTAAGCTAAGTTCGAGAGGTGAAGTACGACGAAATA 720  
QY 721 TCATGGTTTATTAATATGATTTTAAATTAATCTTCAATAAATTTCTTTAGGAAACA 780  
DB 721 TCATGGTTTATTAATATGATTTTAAATTAATCTTCAATAAATTTCTTTAGGAAACA 780  
QY 781 TATCATTTAATGTTTGAAGAACGTGCACATAAGAAAACCTAAGAACGATGAGTTGGAAA 840  
DB 781 TATCATTTAATGTTTGAAGAACGTGCACATAAGAAAACCTAAGAACGATGAGTTGGAAA 840  
QY 841 CAAGAGAAAACACAGCCTTAAGCCTTCTGATCCTCTAGTTGGAGGTTGATTTTCAAC 900  
DB 841 CAAGAGAAAACACAGCCTTAAAGCCTTCTGATCCTCTAGTTGGAGGTTGATTTTCAAC 900  
QY 901 GCATGATAAACGAGAAAGCTCATTAGCACATTTACTTAGATATTTATAATTTATAA 960  
DB 901 GCATGATAAACGAGAAAGCTCATTAGCACATTTACTTAGATATTTATAATTTATAA 960  
QY 961 TGAAAAAATTTTATTTGAAATTTTAAACAATGATGATCAATAAATTTTAAAC 1020  
DB 961 TGAAAAAATTTTATTTGAAATTTTAAACAATGATGATCAATAAATTTTAAAC 1020  
QY 1021 ACACCAATTTAAACCTTTAAAGCATCTTAATAGGAAACGAGAAAGTTAAAGATTCA 1080  
DB 1021 ACACCAATTTAAACCTTTAAAGCATCTTAATAGGAAACGAGAAAGTTAAAGATTCA 1080  
QY 1081 GAAGTGTGTAATGAATAAATGAGGAGGAGAAATGAATGTTAGAGTTTAAATGTCCTTTT 1140  
DB 1081 GAAGTGTGTAATGAATAAATGAGGAGGAGAAATGAATGTTAGAGTTTAAATGTCCTTTT 1140  
QY 1141 GATTAATATTTAAATGAAGAGGAGGAGAAATGAATGTTAGAGTTTAAATGTCCTTTT 1200  
DB 1141 GATTAATATTTAAATGAAGAGGAGGAGAAATGAATGTTAGAGTTTAAATGTCCTTTT 1200  
QY 1201 GTGGTAGAAAAATTTATTTGCGATACACTCCCGAGAGCGGTGCTGCTGCTGGGCGA 1260  
DB 1201 GTGGTAGAAAAATTTATTTGCGATACACTCCCGAGAGCGGTGCTGCTGCTGGGCGA 1260  
QY 1261 AGCGTCTTTTCTGTTGGAAGAAAACCTGCTTAAAGAGGAAACAGAAAGAGCCAGCTTT 1320  
DB 1261 AGCGTCTTTTCTGTTGGAAGAAAACCTGCTTAAAGAGGAAACAGAAAGAGCCAGCTTT 1320  
QY 1321 GTTGTCAACCGTCTCACCAGAAACGAAACAAAAGCCCCCACCCTAAACCTCCTCGATCC 1380  
DB 1321 GTTGTCAACCGTCTCACCAGAAACGAAACAAAAGCCCCCACCCTAAACCTCCTCGATCC 1380  
QY 1381 GACCGAGACTCCTCATTTTCAGCGCGCAGCGGAGAGCAGCGAGGTTTCGTCGAC 1440  
DB 1381 GACCGAGACTCCTCATTTTCAGCGCGCAGCGGAGAGCAGCGAGGTTTCGTCGAC 1440

RESULT 3

AAA07481

ID AAA07481 standard; DNA; 4341 BP.

XX AAA07481;

XX 17-JUL-2000 (first entry)

XX Rice adenylate kinase DNA sequence fragment #1.

XX Rice; adenylate kinase; promoter; structural gene expression;  
XX plant gene expression; herbicide resistance; stress resistance;  
XX light resistance; antibacterial production; flowering induction;  
XX insecticidal substance production; ss.

OS Oryza sativa.

XX WO200015812-A1.

XX



KW insecticide; light resistance; sucrose induction; flower promotion;  
KW increase crop yield; herbicide resistance; ds.  
XX  
OS Oryza sativa.  
XX  
XX WO200015811-A1.  
PN  
XX  
PD 23-MAR-2000.  
XX  
PF 10-SEP-1998; 98WO-JP04088.  
XX  
XX  
PR 10-SEP-1998; 98WO-JP04088.  
XX  
PA (NISC ) NISSAN CHEM IND LTD.  
XX  
XX Uchimiya H, Arai S, Fushimi T, Tagawa M;  
PI  
XX  
XX WPI; 2000-271447/23.  
XX  
XX Rice adenylate kinase gene-originated DNA fragment with promoter  
PT function in plants for controlling expression of desired foreign  
PT structural genes, useful for herbicide resistance and in sucrose  
PT induction in recombinant plants -  
XX  
XX Claim 1; Page 23-25; 30pp; Japanese.

CC This sequence represents a rice adenylate kinase gene-originated DNA  
CC fragment with promoter function in plants. The fragment can be used for  
CC controlling the expression of desired foreign structural genes. The DNA  
CC fragment is used to create a vector, which in turn is used to transform  
CC bacteria or plant cells. The DNA fragment can be inserted into a plant to  
CC control the expression of desired foreign structural genes, useful for  
CC inducing in the plant e.g. herbicide, stress and light resistance.  
CC production of antibacterial and insecticidal substances, sucrose  
CC induction in recombinant plants, and in promoting flowering and crop  
CC yield.  
XX

SQ Sequence 4341 BP; 1284 A; 758 C; 880 G; 1419 T; 0 other;

Query Match 100.0%; Score 1440; DB 21; Length 4341;  
Best Local Similarity 100.0%; Pred. No. 1.2e-287;  
Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CTGCGAGAGATTAATAGTGGGACACACACCAACCCCTGTGGTGGTGGACGCCCTGTGTT 60
DB 1 CTGCGAGGAGATTAATAGTGGGACACACACCAACCCCTGTGGTGGTGGACGCCCTGTGTT 60
QY 61 AATCAACTGGGCTGTTGGTGGACATGGTGTGTCAGGAAAATTAAGCAAGAAAATTAAG 120
DB 61 AATCAACTGGGCTGTTGGTGGACATGGTGTGTCAGGAAAATTAAGCAAGAAAATTAAG 120
QY 121 AAGATGCTCAAGCTGACATGAGAAAACGTAATCCAATGGAAGCAATTTCAAGTCGTTTC 180
DB 121 AAGAATGCTCAAGCTGACATGAGAAAACGTAATCCAATGGAAGCAATTTCAAGTCGTTTC 180
QY 181 TCTGTACTACCATGTTTAGAATACATAGACAGTGCACAGCTGTTGATGGCTCCTATTGG 240
DB 181 TCTGTACTACCATGTTTAGAATACATAGACAGTGCACAGCTGTTGATGGCTCCTATTGG 240
QY 241 CTCGTGTGATACGACTGTGTGCACAAAGCATCAAAATGCTTCCTGGAGTATCTTTATTA 300
DB 241 CTCGTGTGATACGACTGTGTGCACAAAGCATCAAAATGCTTCCTGGAGTATCTTTATTA 300
QY 301 CCGAAAACCCCAAGATTAATCTATTTCACCTCAGGGTAATGTCGTAACATGCAATGCAATG 360
DB 301 CCGAAAACCCCAAGATTAATCTATTTCACCTCAGGGTAATGTCGTAACATGCAATGCAATG 360
QY 361 AATACAAATTCGCAAAATATCATGTTATCTATCTTGTCTCAAAATGAAATTTGAGTCCAA 420
DB 361 AATACAAATTCGCAAAATATCATGTTATCTATCTTGTCTCAAAATGAAATTTGAGTCCAA 420
QY 421 CTGAGACTGCAATACGATTTTCTTTTCAAAAAGAAAATTAATTAATTTTTCATGAAA 480
```

```
DB 421 CTGAGACTGCAATACGATTTTCTTTTCAAAAAGAAAATTAATTAATTTTTCATGAAA 480
QY 481 CGCAATTCAACCGTTTCGAGAAAATATGCTGTCAATAAATAGTAGTCTAGTCAGAAAACAAA 540
DB 481 CGCAATTCAACCGTTTCGAGAAAATATGCTGTCAATAAATAGTAGTCTAGTCAGAAAACAAA 540
QY 541 ATTAATATCAATAAAAAGAGGTTGTTAATTAACAAACCATGTTTCGTACTCAACTCT 600
DB 541 ATTAATATCAATAAAAAGAGGTTGTTAATTAACAAACCATGTTTCGTACTCAACTCT 600
QY 601 AATTTGTAATTTCTTATTTTCAGTCACAAAATTCCAATTTCCAATTAAGAAAATAAAGCT 660
DB 601 AATTTGTAATTTCTTATTTTCAGTCACAAAATTCCAATTTCCAATTAAGAAAATAAAGCT 660
QY 661 AGACGGCTAAGCCCAACCCATCTAAGGCTAAGTTCGAGAGGTGAAGTACGCACGAAAATA 720
DB 661 AGACGGCTAAGCCCAACCCATCTAAGGCTAAGTTCGAGAGGTGAAGTACGCACGAAAATA 720
QY 721 TGATGTTTATTAATATGATTTTAAATAAATTTTAAATAAATTTTACATATAATTTCTTAGGAAACA 780
DB 721 TGATGTTTATTAATATGATTTTAAATAAATTTTAAATAAATTTTACATATAATTTCTTAGGAAACA 780
QY 781 TATCATTTAATGTTTGAATAACCGTCACATTAAGAAAACCTAAGAAAACGATGAGTTGGAAA 840
DB 781 TATCATTTAATGTTTGAATAACCGTCACATTAAGAAAACCTAAGAAAACGATGAGTTGGAAA 840
QY 841 CAAGAGAAAACACAGCCCTTAAGGCTTCTTGATCCTCTAGTTCGAGGTTGATTTTCAAC 900
DB 841 CAAGAGAAAACACAGCCCTTAAGGCTTCTTGATCCTCTAGTTCGAGGTTGATTTTCAAC 900
QY 901 GCATGATAAACGAGAAAAGCTCATTAGCACATTAATTAATTAATTTATTAATTAATTA 960
DB 901 GCATGATAAACGAGAAAAGCTCATTAGCACATTAATTAATTAATTTATTAATTAATTA 960
QY 961 TGAAAAAATATTTATTTGAATTTTAAACAATGTATGATATAATTTTAAATTTTAAAC 1020
DB 961 TGAAAAAATATTTATTTGAATTTTAAACAATGTATGATATAATTTTAAATTTTAAAC 1020
QY 1021 ACACCAATTTAACCCCTTTAAAAAGCATCTTAATAGGAAAACGAGGAAGTTAAAGATTCAAC 1080
DB 1021 ACACCAATTTAACCCCTTTAAAAAGCATCTTAATAGGAAAACGAGGAAGTTAAAGATTCAAC 1080
QY 1081 GAGGTTTGGATTAATGAAAATGGGTTGGATTAAGATTTGGTAAATGAAATCAGGTTAG 1140
DB 1081 GAGGTTTGGATTAATGAAAATGGGTTGGATTAAGATTTGGTAAATGAAATCAGGTTAG 1140
QY 1141 GATTAAATATTAATAATGAAGAGGAGAAATGAATGGTTAGAGTTTAAATGCTCTTTTG 1200
DB 1141 GATTAAATATTAATAATGAAGAGGAGAAATGAATGGTTAGAGTTTAAATGCTCTTTTG 1200
QY 1201 GTGGTAGAAAATTTATTTGCCATACACTCCCGAGAGCGGTGCTGCTGCGTGGGAGA 1260
DB 1201 GTGGTAGAAAATTTATTTGCCATACACTCCCGAGAGCGGTGCTGCTGCGTGGGAGA 1260
QY 1261 AGCGTCTTTTTCGTTGAAAAAAAACCTGTTTAAAAAGGAAAACAGAAAAGAGCCAGCTTT 1320
DB 1261 AGCGTCTTTTTCGTTGAAAAAAAACCTGTTTAAAAAGGAAAACAGAAAAGAGCCAGCTTT 1320
QY 1321 GTTGTACCGCTCTCACGAGAAAACGAAAACAAAAGCCCCACCTTAAACCTCTCGATCC 1380
DB 1321 GTTGTACCGCTCTCACGAGAAAACGAAAACAAAAGCCCCACCTTAAACCTCTCGATCC 1380
QY 1381 GACCGAGACTCCTCCATTTTCAGCGGCGCACGCGGAGAGCACGCGGAGAGTTCGTCGAC 1440
DB 1381 GACCGAGACTCCTCCATTTTCAGCGGCGCACGCGGAGAGCACGCGGAGAGTTCGTCGAC 1440
```

RESULT 5  
ABL49953  
ID ABL49953 standard; DNA; 422 Bp.  
XX  
AC ABL49953;  
XX

```

DT 10-JUN-2002 (first entry)
DE Rice Wanderer-Osl nucleotide sequence SEQ ID NO:28.
XX
XX Rice; plant; detection; polymorphism; transposable element; genome;
KW gene; ds.
XX
OS Oryza sativa.
PN WO200212484-A1.
XX
XX 14-FEB-2002.
XX
XX 02-AUG-2001; 2001WO-JP06661.
XX
XX 02-AUG-2000; 2000JP-0234577.
XX
XX (NISR) JAPAN TOBACCO INC.
PA (SYGN) SYNGENTA LTD.
XX
XX Komori T, Nitta N;
PI WPI; 2002-241759/29.
XX
XX Construction of marker for detecting plant genome polymorphism with use
PT of transposable element, useful particularly in studying restriction
PT fragment length polymorphism applicable in cell genetics -
XX
XX Example 2; Fig 7; 110pp; Japanese.
XX
XX The present invention describes a method for constructing a marker for
CC detecting polymorphisms in a plant genome. The method comprises the
CC production of a primer for nucleic acid amplification by using the base
CC sequences of a transposable element and/or the domain adjacent to it.
CC Also described is a marker for detection polymorphisms in plant genomes.
CC The constructed marker can be used for detecting plant genome
CC polymorphisms, which is useful particularly in studying restriction
CC fragment length polymorphism applicable in cell genetics e.g. for
CC analysing and selecting specific breeds of plants. The method is simple,
CC easy, less time consuming and not so laborious, e.g. in the study of
CC less frequently occurring polymorphism between various species to enable
CC identification of the site and isolation of the required gene.
CC ABL49926 to ABL50036 represent nucleotide sequences used in the
CC exemplification of the present invention.
XX
XX Sequence 422 BP; 145 A; 65 C; 56 G; 156 T; 0 other;
SQ
Query Match 6.9%; Score 100; DB 24; Length 422;
Best Local Similarity 70.5%; Pred No. 1,9e-11;
Matches 148; Conservative 0; Mismatches 60; Indels 2; Gaps 1;
QY 869 TTGATCCTCTAGTTGGAGTTGATTTTCAACGCGATGATAAACGAGAAAGCTCATTAGCA 928
DB 82 TTAATCTCCACATTTGGAGATAAATTTTGGTTCACGTAATAAACGAGAAAGCTCATTAGCA 141
QY 929 CATTTACTTAGATTTTATATTAATTAATTAACCTTGAAAAAATATTTTGAATTTTAA 988
DB 142 CATATTAATTAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTCTTTTAA 201
QY 989 AACAA--TGATGCTAAATTTATTTTAAACACACAAATTAACCTTTTAAAGCA 1046
DB 202 AACAACTTCTATATAGAACTTTTAAACAAAAATGTATCATTTTAAACAAATTTGAAAGCG 261
QY 1047 TCTTAATAGAAACGAGGAAGTTAAAGATT 1076
DB 262 TGCTAACGAAACAAAGGAAGTTGAGAAAT 291
RESULT 6
ABL49955
ID ABL49955 standard; DNA; 423 BP.
XX
AC ABL49955;

```

```

XX 10-JUN-2002 (first entry)
XX
XX Rice Wanderer-Osl nucleotide sequence SEQ ID NO:30.
DE
XX
XX Rice; plant; detection; polymorphism; transposable element; genome;
KW gene; ds.
XX
XX Oryza sativa.
XX
XX WO200212484-A1.
XX
XX 14-FEB-2002.
XX
XX 02-AUG-2001; 2001WO-JP06661.
XX
XX 02-AUG-2000; 2000JP-0234577.
XX
XX (NISR) JAPAN TOBACCO INC.
PA (SYGN) SYNGENTA LTD.
XX
XX Komori T, Nitta N;
PI WPI; 2002-241759/29.
XX
XX Construction of marker for detecting plant genome polymorphism with use
PT of transposable element, useful particularly in studying restriction
PT fragment length polymorphism applicable in cell genetics -
XX
XX Example 2; Fig 7; 110pp; Japanese.
XX
XX The present invention describes a method for constructing a marker for
CC detecting polymorphisms in a plant genome. The method comprises the
CC production of a primer for nucleic acid amplification by using the base
CC sequences of a transposable element and/or the domain adjacent to it.
CC Also described is a marker for detection polymorphisms in plant genomes.
CC The constructed marker can be used for detecting plant genome
CC polymorphisms, which is useful particularly in studying restriction
CC fragment length polymorphism applicable in cell genetics e.g. for
CC analysing and selecting specific breeds of plants. The method is simple,
CC easy, less time consuming and not so laborious, e.g. in the study of
CC less frequently occurring polymorphism between various species to enable
CC identification of the site and isolation of the required gene.
CC ABL49926 to ABL50036 represent nucleotide sequences used in the
CC exemplification of the present invention.
XX
XX Sequence 423 BP; 146 A; 66 C; 56 G; 155 T; 0 other;
SQ
Query Match 6.8%; Score 97.4; DB 24; Length 423;
Best Local Similarity 69.7%; Pred No. 6.4e-11;
Matches 147; Conservative 0; Mismatches 61; Indels 3; Gaps 1;
QY 869 TTGATCCTCTAGTTGGAGTTGATTTTCAACGCGATGATAAACGAGAAAGCTCATTAGCA 928
DB 82 TTAATCTCCACATTTGGAGATAAATTTTGGTTCACGTAATAAACGAGAAAGCTCATTAGCA 141
QY 929 CATTTACTTAGATTTTATATTAATTAATTAACCTTGAAAAAATATTTTGAATTTTAA 988
DB 142 CATATTAATTAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTTCTTTTAA 201
QY 989 AACAA--TGATGCTAAATTTATTTTAAACACACA---CCAATTTAACCTTTTAAAGCA 1045
DB 202 AACAACTTCTATATAGAACTTTTAAACAAAAATGTATCATTTTAAACAAATTTGAAAGC 261
QY 1046 ATCCTAATAGGAAACGAGGAAGTTAAAGATT 1076
DB 262 TGCTAACGAAACAAAGGAAGTTGAGAAAT 292
RESULT 7
ABL49954
ID ABL49954 standard; DNA; 424 BP.
XX

```





```
XX PA (EPIG-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K;
XX XX WPI; 2002-130909/17.
XX XX Nucleic acid comprising fragment of chemically modified gene, useful
XX PT for diagnosis and treatment of diseases associated with abnormal
XX PT cytosine methylation
XX XX Claim 1; SEQ ID NO 330; 32pp + Sequence Listing; German.
XX XX
XX CC The present invention provides a number of human immune system associated
XX CC genes which are modified by the methylation of cytosines. The sequences
XX CC can be used in the diagnosis and treatment of immune system disorders,
XX CC including eye diseases such as retinopathy, neovascular glaucoma and
XX CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
XX CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
XX CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
XX CC diseases. The present sequence is a gene of the invention.
XX XX
XX SQ Sequence 9642 BP; 2381 A; 88 C; 2349 G; 4824 T; 0 other;

Query Match 4.4%; Score 64; DB 24; Length 9642;
Best Local Similarity 45.8%; Pred. No. 0.00091;
Matches 307; Conservative 0; Mismatches 350; Indels 13; Gaps 2;

QY 403 ATTGAATTTGAGTCCAACTGAGACTGCAATACGATTTTCTTTTCAAAAAGAAATATT 462
DB 4113 ATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4054
QY 463 AATTTTTTTTTTTCATGAAGCAATTCACCGTTCGAGAAATATCGTGCATAAATAAGTA 522
DB 4053 AATTAATTTTCATATACCAACTAATATACATATATATATATATATATATATATAT 3994
QY 523 GTCTAGTCGAGAAACAAATTAATATACATAAATAAAGAGGTTGTTAATTAACAACCAT 582
DB 3993 CTCCTAATTAACCTATCCAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3934
QY 583 GTTTCGTACTACAACCTAATTTGTAATTTCTTATTTTCAGTCACAAAATTCCAATTCCA 642
DB 3933 CCT-----AAAAAATTAACCCATAAATTTACTAATAATCACACAATACTATATAA 3880
QY 643 ATTAAGAAAAATAAAGTAGAGCGCTAAGCCACCCATCTAAGGCTAAGTTCGAGAGGTG 702
DB 3879 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3820
QY 703 AAGTAGCGACGAAATATGATGGTTTATTAATATGATTTTTTTTTTAATAACTTTTCACAT 762
DB 3819 AAAAAATTCCTCAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 3760
QY 763 AATTTCTTTAGGAACATATCATTTAATGTTTGAATAACGTCGACATAGAAAACATAA 822
DB 3759 CACTACCTTAAATAACCAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3700
QY 823 GAACATGATGTTGGGAACAAGAGAAAAACACAGCCTTTTAAGGCTTTCTGTATCCTAGTT 882
DB 3699 AAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 3640
QY 883 GGAGGTGATTTTCAACGCAATGATAAACGAGAAAGCTCAITAGCACATTATTACTTTAGA 942
DB 3639 ATAACAAAAATATACATATAATAATTTAAACACAATTTCAACAAAAATATTATAATAA 3580
QY 943 TATTTATTAATTAATTAATTAATTAATTAATTTTATTTGATTTTAAACAATGATGCAT 1002
DB 3579 ACCAATAAATAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3527
QY 1003 AATATTATTTTAAACACACCAATTTAAACCTTTTAAAGAGCTCTCTAATAGCAACGA 1062
DB 3526 AACTATTCTTTAAACACATTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3467
QY 1063 GGAAGTTAAA 1072
```

```
Db 3466 AAAAAATAAA 3457
RESULT 11
AAS46310/C
ID AAS46310 standard; DNA; 6095 BP.
XX AC AAS46310;
XX XX 18-DEC-2001 (first entry)
XX DT Tumour suppressor gene derived chemically modified sequence #32.
XX DE Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
XX KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
XX KW cytosine methylation; ds.
XX OS Homo sapiens.
XX PN WO200168912-A2.
XX PD 20-SEP-2001.
XX PF 15-MAR-2001; 2001WO-EP02955.
XX PR 15-MAR-2000; 2000DE-1013847.
XX PR 06-APR-2000; 2000DE-1019058.
XX PR 07-APR-2000; 2000DE-1019173.
XX PR 30-JUN-2000; 2000DE-1032529.
XX PR 01-SEP-2000; 2000DE-1043826.
XX PA (EPIG-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K;
XX DR WPI; 2001-602752/68.
XX PT Fragments of chemically modified genes associated with tumour suppressor
XX PT genes and oncogenes, useful in designing primers and probes for
XX PT analysing diseases associated with cytosine methylation state e.g.
XX PT cancer
XX PS Claim 1; SEQ ID No 32; 27pp; English.
XX CC The invention relates to a nucleic acid comprising a sequence of 18
XX CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
XX CC bisulphite, of genes associated with tumour suppression and
XX CC oncogenes having a sequence taken from 536 (actually 533 since
XX CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
XX CC (SS) and sequences complementary to (SS). The nucleic acid may be a
XX CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
XX CC form part of a set of probes for detecting the cytosine methylation state
XX CC and/or single nucleotide polymorphisms and also to be used in an
XX CC array for analysing diseases associated with CpG dinucleotides e.g.
XX CC cancers and tumours. The probes can also be used in a method for
XX CC ascertaining genetic and/or epigenetic parameters for the diagnosis
XX CC and/or therapy of existing diseases or the predisposition to specific
XX CC diseases, by analysing cytosine methylations. The parameters may be
XX CC compared to another set of genetic and/or epigenetic parameters, the
XX CC differences serving as basis for diagnosis and/or prognosis events which
XX CC are disadvantageous to patients. The present sequence is one of the
XX CC 533 genomic sequences derived from tumour suppressor genes and
XX CC oncogenes. Sequences with even numbered Seq ID numbers are the
XX CC complementary sequence of the corresponding odd numbered sequence (e.g.
XX CC ID 2 and ID1, ID 536 and ID 535, except for those whose partner sequence
XX CC is missing).
XX CC Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic
XX CC format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 6095 BP; 1666 A; 108 C; 1392 G; 2929 T; 0 other;
```



```
Query Match      4.4%; Score 63; DB 22; Length 6095;
Best Local Similarity 44.6%; Pred. No. 0.0013;
Matches 287; Conservative 0; Mismatches 355; Indels 1; Gaps 1;

QY 461 TTAATTTTTCATGAAACGCAATCAACCGTTTCGAGAAATATGCTGTATATAAATAAG 520
DB 2571 TTACTTTTAAATTAATTCACATAACAATAAATACTACGTTTAAATTTCCAAACATTAACCA 2512
QY 521 TAGTCTAGTCGAGAACCAAAATTAATATACATATAAAGAGGCTTCTTAATTACAAACC 580
DB 2511 TAATATATTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2452
QY 581 ATGTTTCGTAACACTCTAATTTGTAATTTCTTATTTTCAATATAAATAAATAAATAAATA 640
DB 2451 ACTTCCCTTAAATAATTTACAAACTCTTCTCTCTTTTAAATCTCTAAATAAATAAATAA 2392
QY 641 CAAATTAAGAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 700
DB 2391 TAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2332
QY 701 TGAAGTACGACGAAATAATATGATGTTTATTAATATGATGTTTATTAATATGATGTTTAT 760
DB 2451 ACTTCCCTTAAATAATTTACAAACTCTTCTCTCTTTTAAATCTCTAAATAAATAAATAA 2392
QY 641 CAAATTAAGAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 700
DB 2391 TAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2332
QY 701 TGAAGTACGACGAAATAATATGATGTTTATTAATATGATGTTTATTAATATGATGTTTAT 760
DB 2331 TAAACGTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2272
QY 761 ATAAATTTCTTTAGGAACATATCAATTTAATGTTTGAACACGTCACATACAAACT 820
DB 2271 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2212
QY 821 AAGAAGTATGAGTTGGAAACAGAGAAACACAGCGCTTAAGGCTTCTTGTATCCTCTAG 880
DB 2211 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2152
QY 881 GAGGAAGTTAAGATTCACCGAAGTGTGGAATATGAAAAAT 1103
DB 1972 TAAACGATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1930

RESULT 12
ABL70150/c
ID ABL70150 standard; DNA; 6095 BP.
XX AC ABL70150;
XX DT 01-JUL-2002 (first entry)
XX DE Chemically treated cell signalling DNA sequence complementary to#20.
XX KW Cell signalling; cytosine methylation; cell signalling disease;
XX KW cancer; tumour; cytosine; ds.
XX OS Unidentified.
XX PN WO200202807-A2.
XX PD 10-JAN-2002.
XX PF 29-JUN-2001; 2001WO-EP07471.
XX PR 30-JUN-2000; 2000DE-1032529.
XX PR 01-SEP-2000; 2000DE-1043826.
```

(EPIG-) EPIGENOMICS AG.  
Olek A, Piepenbrock C, Berlin K,  
WPI; 2002-154758/20.

Nucleic acid, useful for diagnosis and therapy of diseases associated with cell signalling e.g. cancer, comprises chemically modified genomic sequences of genes associated with cell signalling.

Claim 1; SEQ ID NO 40; 24pp+sequence listing; English.

The invention relates to a nucleic acid comprising a sequence of at least 18 bases of a segment of chemically pretreated DNA of genes associated with cell signalling. The activity of the modified sequences of the invention may be described as cytostatic. The object of the invention is to provide the chemically modified DNA of genes associated with cell signalling, as well as oligonucleotides and/or PNA-oligoners for detecting cytosine methylation, as well as a method which is particularly suitable for the diagnosis and/or therapy of genetic and epigenetic parameters of genes associated with cell signalling. The chemically modified DNA provided by the invention is useful for diagnosis and therapy of diseases such as solid tumours and cancer. The sequences given in records ABL70111-ABL70626 represent chemically pre-treated genomic DNA's of genes associated with cell signalling.

Note: The sequence data for this patent is not represented in the printed specification, but is based on sequence information supplied by the European Patent Office.

Sequence 6095 BP; 1666 A; 108 C; 1392 G; 2929 T; 0 other;

Query Match 4.4%; Score 63; DB 24; Length 6095;  
Best Local Similarity 44.6%; Pred. No. 0.0013;  
Matches 287; Conservative 0; Mismatches 355; Indels 1; Gaps 1;

QY 461 TTAATTTTTCATGAAACGCAATCAACCGTTTCGAGAAATATGCTGTATATAAATAAG 520  
DB 2571 TTACTTTTAAATTAATTCACATAACAATAAATACTACGTTTAAATTTCCAAACATTAACCA 2512  
QY 521 TAGTCTAGTCGAGAACCAAAATTAATATACATATAAAGAGGTTGTTAAATACAAACC 580  
DB 2511 TAATATATTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2452  
QY 581 ATGTTTCGTAACACTCTAATTTGTAATTTCTTATTTTCAATATAAATAAATAAATAAATA 640  
DB 2451 ACTTCCCTTAAATAATTTACAAACTCTTCTCTCTTTTAAATCTCTAAATAAATAAATAA 2392  
QY 641 CAAATTAAGAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 700  
DB 2391 TAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2332  
QY 701 TGAAGTACGACGAAATAATATGATGTTTATTAATATGATGTTTATTAATATGATGTTTAT 760  
DB 2331 TAAACGTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2272  
QY 761 ATAAATTTCTTTAGGAACATATCAATTTAATGTTTGAACACGTCACATACAAACT 820  
DB 2271 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2212  
QY 821 AAGAAGTATGAGTTGGAAACAGAGAAACACAGCGCTTAAGGCTTCTTGTATCCTCTAG 880  
DB 2211 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2152  
QY 881 GAGGAAGTTAAGATTCACCGAAGTGTGGAATATGAAAAAT 1103  
DB 1972 TAAACGATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1930





```
PS Claim 1; SEQ ID NO 28; 23pp + Sequence Listing; English.
XX
CC The present invention provides a number of human metastasis associated
CC genes which are modified by cytosine methylation. The sequences can be
CC used in the diagnosis and treatment of cancer. The present sequence is
CC one of the genes of the invention.
XX
SQ Sequence 6095 BP; 1666 A; 108 C; 1392 G; 2929 T; 0 other;
  Query Match      4.4%; Score 63; DB 24; Length 6095;
  Best Local Similarity 44.6%; Pred. No. 0.0013;
  Matches 287; Conservative 0; Mismatches 355; Indels 1; Gaps 1;
  QY 461 TTAATTTTTCATGAACGCAATTCACCGTTCGAGAAATATGCTGCATATAATAG 520
  DB 2571 TTACTTTTAAATTCATACAAATATAATTAATTAATTTTCCAAACATAAACCA 2512
  QY 521 TAGTCTAGTCAGAAACAAATTAATATACATAAAAAAGAGGTTGTTAATTACAAACC 580
  DB 2511 TAATATATTATAAATAATAAATTAATTAATTAATTAATTAATTAATTAATTAAT 2452
  QY 581 ATGTTTCGTACTACAATCTTAATTTGTAATTCCTATTTCAGTCACAAAATTCCAATTC 640
  DB 2451 ACTTCCCTCAAAAATATTACAAACTCTCTCTCTTTAAAAATCTACTAAATAAAAAA 2392
  QY 641 CAATTAAGAAAAATAACGTAGCGCTTAAGCCACCATCTAAGGCTTAAGTTTCGAGAG 700
  DB 2391 TAAAAAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2332
  QY 701 TGAAGTACGCACGAAAAATATGATGGTTTATTAATATGATTTTTTTTAAATACCTTCAC 760
  DB 2331 TAAACGTATAAATAAATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2272
  QY 761 ATAAATTTCTTGAACACATATCATTTAATGTTTGGTTTGAACACGTGCACATAGAAACT 820
  DB 2271 AAATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2212
  QY 821 AAGAACGATGAGTTGGGAAACAGAGAAAAACACAGCGCTTAAAGGCTTCTTGGATCCTAG 880
  DB 2211 AATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2152
  QY 881 TTGAGGTTGTTTTCACACGATGATAAGGAGAAAGCTCATTAGCACATTAATTACTTA 940
  DB 2151 AAAAAATATATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2093
  QY 941 GATATTTATATAAATTAACCTTGAAAAAATAATTTATTTGTAATTTTTTAAACAATGTATGC 1000
  DB 2092 TAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2033
  QY 1001 ATAAATTTATTTTAAACACACCAATTTTAACCTTTTAAAGCATCCTTAATAGGAAC 1060
  DB 2032 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1973
  QY 1061 GAGGAAGTTAAAGATTCCACGAGTGTGTTGATAATGAAAAAT 1103
  DB 1972 TAAACGATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1930
  RESULT 15
  AAS46802/c
  ID AAS46802 standard; DNA; 8237 BP.
  XX
  AC AAS46802;
  XX
  XX Tumour suppressor gene derived chemically modified sequence #528.
  DT 18-DEC-2001 (first entry)
  XX
  DE Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
  XX cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
  KW cytosine methylation; ds.
  XX
  OS Homo sapiens.
```

Db 4082 CCATCTCAAAACAATTATTTCACCTTCACTTAAAAATTAAAAATATATACTCTAACCAATACA 4023  
Qy 680 TCTAAGGCTTAAGTTTCGAGAGGTGA-AGTACGCACGAAAAATATGATGGTTTTATTATATG 738  
Db 4022 TCTAAATTTATTAT 3963  
Qy 739 ATTTTAAATAAATCTTTCACATAAATTTCTTTAGGAAACATATCATTTAATGGTTGA 798  
Db 3962 AATCTAATCGTCTAAACCCCTTTTCCAAATAAATAAAAAAATTTAATTTCAATAATTATCA 3903  
Qy 799 AAAACGTGCACATAAGAAAACTAAGAACGATGAGTTGGGAAACAAGAAAAACACAGCC 858  
Db 3902 CTATTAAATTTTATCTAACCTTTAAACTTAATTAATTCATCCCACTATATTACACTAATAT 3843  
Qy 859 TTAAGGCTTCTTGAPCCTCTAGTTGGAGGTTGATTTTCAACGCATGATAAACGAGAAAG 918  
Db 3842 ATAAAACTCCCAAAACGTAAAAATAAACTCATCTATCCTTTACGAATAAATATCAAAATTA 3783  
Qy 919 CTCATTAGCACATTTATTACTTAGATATTATATATATATATATATATATATATATATAT 978  
Db 3782 ACTTACTTCATAAATTAATATATTTTTCATTAATAATTTTAAATAATTTTAAATAATATATA 3723  
Qy 979 GAATTTTAAACAATGTATGCATAAATTTATTTTAAACACACCAATTTAACCCCTT 1038  
Db 3722 ATATAATTTAAAAATTTTAAAAAAATAATATATTTATACAAATACACCTATTAAAAACCTT 3663  
Qy 1039 AAAAGCATCCTTAATAGGAACGAGGATTTAAAGATTTCAC 1079  
Db 3662 TAAATTTATATTAAAAATTTAAAAAAATTTAAAAACAATAAC 3622

Search completed: March 15, 2003, 10:56:21  
Job time : 329.521 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2003, 10:25:55 ; Search time 45.5838 Seconds  
(without alignments)  
9687.969 Million cell updates/sec

Title: US-09-802-937-2  
Perfect score: 1440  
Sequence: 1 ctgcaggaagattaattagg.....cgcgacgcaggtctgcgac 1440

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- 1: /cgn2\_6/ptodata/2/ina/5A COMB.seq.\*
- 2: /cgn2\_6/ptodata/2/ina/5B COMB.seq.\*
- 3: /cgn2\_6/ptodata/2/ina/6A COMB.seq.\*
- 4: /cgn2\_6/ptodata/2/ina/6B COMB.seq.\*
- 5: /cgn2\_6/ptodata/2/ina/PTUS COMB.seq.\*
- 6: /cgn2\_6/ptodata/2/ina/backfilesi.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	57.2	4.0	7218	1 US-08-232-463-14	Sequence 14, Appl
C 2	50.2	3.5	6124	4 US-08-213-419B-3	Sequence 3, Appl
C 3	50	3.5	19124	2 US-08-487-826B-13	Sequence 13, Appl
C 4	46.8	3.2	636	4 US-08-998-416-1137	Sequence 1137, Ap
C 5	46.6	3.2	1850	3 US-08-617-860B-32	Sequence 22, Appl
C 6	46.6	3.2	2389	1 US-08-153-583-2	Sequence 2, Appl
C 7	46.6	3.2	2389	2 US-08-460-507-2	Sequence 2, Appl
C 8	46.6	3.2	4098	2 US-08-605-106-4	Sequence 4, Appl
C 9	46.6	3.2	4276	1 US-07-973-324A-3	Sequence 3, Appl
C 10	46.6	3.2	4276	4 US-08-343-380-3	Sequence 3, Appl
C 11	46.6	3.2	4276	4 US-09-072-435-3	Sequence 3, Appl
C 12	46.6	3.2	4276	4 US-09-072-917A-3	Sequence 3, Appl
C 13	45.8	3.2	5852	1 US-07-867-106-2	Sequence 2, Appl
C 14	45.8	3.2	9048	3 US-08-973-273-4	Sequence 4, Appl
C 15	45	3.1	636	4 US-08-998-416-1137	Sequence 1137, Ap
C 16	44.2	3.1	837	4 US-08-998-416-288	Sequence 288, App
C 17	44	3.1	658	4 US-08-998-416-595	Sequence 595, App
C 18	43.8	3.0	3950	4 US-09-398-395A-33	Sequence 33, Appl
C 19	43	3.0	1857	4 US-09-299-378-3	Sequence 3, Appl
C 20	42.8	3.0	6265	4 US-09-129-112-3	Sequence 3, Appl
C 21	42.6	3.0	20674	4 US-09-641-638-651	Sequence 651 App
C 22	42.6	3.0	28001	4 US-09-813-993-3	Sequence 3, Appl
C 23	42.4	2.9	665	2 US-08-883-795A-36	Sequence 36, Appl
C 24	42	2.9	711	4 US-08-998-416-786	Sequence 786, App
C 25	41.8	2.9	3305	3 US-08-714-918-76	Sequence 76, Appl
C 26	41.8	2.9	3305	4 US-09-265-315-76	Sequence 76, Appl
C 27	41.8	2.9	3305	4 US-09-265-315-76	Sequence 76, Appl

C 28	41.8	2.9	3305	4 US-09-266-417-76	Sequence 76, Appl
C 29	41.6	2.9	1431	4 US-09-316-083-2	Sequence 2, Appl
C 30	41.6	2.9	6152	4 US-08-973-482-1	Sequence 1, Appl
C 31	41.4	2.9	665	2 US-08-883-795A-36	Sequence 36, Appl
C 32	41.4	2.9	2226	4 US-08-924-629C-35	Sequence 35, Appl
C 33	41.2	2.9	1983	4 US-09-134-001C-2202	Sequence 2202, Ap
C 34	41.2	2.9	5852	1 US-07-867-106-2	Sequence 2, Appl
C 35	41.2	2.9	35060	3 US-08-814-095-7	Sequence 7, Appl
C 36	41.2	2.9	84495	4 US-09-797-906-3	Sequence 3, Appl
C 37	41	2.8	592	4 US-09-328-111-39	Sequence 39, Appl
C 38	41	2.8	2061	3 US-09-020-033-1	Sequence 1, Appl
C 39	40.8	2.8	724	4 US-08-998-416-683	Sequence 683, App
C 40	40.2	2.8	3850	3 US-08-329-799-34	Sequence 34, Appl
C 41	40.2	2.8	4285	4 US-09-410-464-1	Sequence 1, Appl
C 42	40.2	2.8	6243	2 US-09-056-075-1	Sequence 1, Appl
C 43	40	2.8	2058	2 US-08-749-391-1	Sequence 1, Appl
C 44	40	2.8	2058	3 US-09-390-200-1	Sequence 1, Appl
C 45	40	2.8	6124	4 US-08-213-419B-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1  
US-08-232-463-14/c  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patencin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)883-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZgpt-F1s  
US-08-232-463-14

Query Match 4.0%; Score 57.2; DB 1; Length 7218;

```
Best Local Similarity 9.5%; Pred. No. 0.00011;
Matches 38; Conservative 197; Mismatches 165; Indels 0; Gaps 0;

QY 955 TAAACTTGAAAAAATATTTATTTGAATTTTAAACAATATGATGCAATAAATATTTT 1014
DB 1515 TCAGTTTCAAAAAACGGCATGTAGGCATCACTGTGTAATACCTATCTATGCAAGTAGTTAA 1456
QY 1015 AAAAAACACACCAATTAACCCCTTTTAAAGAGCATCTTAATAGGAACGAGGAAGTTAAGA 1074
DB 1455 AGAGATAGAGAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1396
QY 1075 TTCACCAAGTGTGGATAATGAATAATGGGTGGGATTAAGATGTAATGAATCAG 1134
DB 1395 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1336
QY 1135 GGTTAGGATTAATATTAATAATGAAGAGGAGATGAATGTTAGAGTTTAAATGTGTC 1194
DB 1335 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1276
QY 1195 TTTTGGTGGTAGAAAAATTTTGCATACACTCCCGAGAGCGGTGCGTGGTGGT 1254
DB 1275 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1216
QY 1255 GGCAGAGCGCTTTTTCGTTGGAATAAAAACTGCTTAAAAAGAAACAGAAAGAGCC 1314
DB 1215 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1156
QY 1315 AGCTTTGGTGCACCTCTCACCAAGAAACAGAAACAAAAAG 1354
DB 1155 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1116

RESULT 2
US-08-213-4198-3/c
; Sequence 3, Application US/082134198
; Patent No. 6333406
; GENERAL INFORMATION:
; APPLICANT: Inselburg, J. et al.
; TITLE OF INVENTION: GENE ENCODING PROTEIN ANTIGENS OF PLASMODIUM FALCIPARUM
; FILE OF INVENTION: AND USE THEREFOR
; FILE REFERENCE: J11-002CNCIP
; CURRENT APPLICATION NUMBER: US/08/213,419B
; PRIOR FILING DATE: 1994-03-14
; PRIOR APPLICATION NUMBER: US 07/870,506
; PRIOR FILING DATE: 1992-04-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 6124
; TYPE: DNA
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2407)..(2439)
; NAME/KEY: CDS
; LOCATION: (2598)..(3404)
; NAME/KEY: CDS
; LOCATION: (3580)..(3720)
; NAME/KEY: CDS
; LOCATION: (3850)..(5835)
US-08-213-4198-3

Query Match 3.5%; Score 50.2; DB 4; Length 6124;
Best Local Similarity 44.9%; Pred. No. 0.0053;
Matches 276; Conservative 0; Mismatches 333; Indels 6; Gaps 2;

QY 431 AATACCAATTTCTTTTCAAAAGAAATTAATTAATTTTTTTCATGAACGCAATTCAA 490
DB 1188 AATAATTTATCTAGTACTTACCATATATATATTTTTTTTATAAGCATTTAATA 1129
QY 491 CCGTTCAGAAATATGCTGCTCAATAATAGTAGTCTAGTCAGAAACAAATTAATCA 550
DB 1128 CATTTTATAGAGGTCTCTTAATAAAAAATTTGTTGGTAATACTACCCCAATCATTTTA 1069
```

```
QY 551 CATAAAAAGAGGTTGTTAAATTACAAACCATGTTTCGTTACTACACTCTAATTTGTAA 610
DB 1068 ATTATTTTCATA---TATATATGTATATATATATATATGACTAAATTTAAAGATGA 1013
QY 611 TTTCTATTTTCAGTCACAAAATTCCAATTTTCCAATTAAGAAAAATAAAGCTAGACGCTAA 670
DB 1012 ATAACATATTAGTTATATAGCTATTACATATATGTACATTTTTCATTTCTCTCTTTAA 953
QY 671 GCCACCCATCTAAGGCTAAGTTTCGAGAGGTGAGTACGCACGAAAAATATGATGTTTA 730
DB 952 ATATATTTAAAAAATATATATATGTTTCATTAACATCAATAATATATATATATATA 893
QY 731 TTAATATGATTTTTTTTTTAAATAAATCTTCA--CATAAATTTCTTTAGGAAACATATCATTT 788
DB 892 TATATATATTTTTTTTTTAAATCTCATGAATCGCTTCCATTTTCGTAATGGAATTT 833
QY 789 AATGGTTTGAAGACGTCACATAAGAAAACTAAGAACGATGAGTTGGAAACAGAGAA 848
DB 832 ATTTTTCATATGATATATATATATATATATATATATATATATATATATATATATA 773
QY 849 AAACACAGCCTTAAGGCTTCTTGATCCTCTAGTTGGAGGTTGATTTTCAAACGCGATGA 908
DB 772 AAAAAGAAATACATAATATAGTTTAACTAAGAACTAGTTTAAATTCATATTTTAAATAT 713
QY 909 AACGAGAAAGCTCATTAGCACATTTACTTTAGATATTTTATATATATATATATATATA 968
DB 712 TAAATGAAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 653
QY 969 ATATTTATTTGAATTTTTTAAACAATGTATGCAATAAATTTTAAAAAACACACCAAT 1028
DB 652 AACAAACTTGTATAAAATTTTCAATATGATTATTCATTTTATATATATTTTCATAAAAT 593
QY 1029 TTAACCCCTTTAAAAA 1043
DB 592 TTTATACATAACAG 578

RESULT 3
US-08-487-826B-13
; Sequence 13, Application US/08487826B
; Patent No. 5993827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,826B
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CPI
; TELECOMMUNICATION INFORMATION:
```

TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19124 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-487-826B-13

Query Match 3.5%; Score 50; DB 2; Length 19124;  
Best Local Similarity 45.0%; Pred. No. 0.0085;  
Matches 273; Conservative 0; Mismatches 325; Indels 8; Gaps 2;

QY 483 CAAATTCACGCGTCGAGAAATATGCTGTCATTAATAAGTAGTCTAGTGCAGAAACAAAT 542  
DB 15593 CAAAAAAATTAATAAAAAATTTTATATAAAAAAAATGATTAAAAAAATAAAAA 15652  
QY 543 TAATATCACATAAAAAAGAGTGTAAATTACAAACCATGTTTCGTAACACTCTAA 602  
DB 15653 CAAAGAAGAAAAAACAATTAAAAAATAAATAATATATATATATATATATATATAT 15712  
QY 603 TTGTAAATCTTATTCAGTCACAAATTCGAATTCGAATTAAGAAATAAAGCTAG 662  
DB 15713 AAGAAAAAATATATTAATAATAAATATATATATATATATATATATATATATATAT 15771  
QY 663 ACGGCTAAGCCCAACCATCTAAGGCTAAGTTCGAGAGTGAAGTACGACGAAAAATATG 722  
DB 15772 ATGTTAAAAAATAAT 15831  
QY 723 ATGTTTAT 782  
DB 15832 TAAATAAAAATTTTAAATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 15891  
QY 783 TCATTTAATGGTTGAAAAACGTGCACATAAGAAAACTAAGAACGATGAGTGGGAACA 842  
DB 15892 AAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 15951  
QY 843 AGAGAAAAACACAGCCTTAAGGCTTCTTGATCTCTAGTGGAGTGTATTTTCNAACGC 902  
DB 15952 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 16004  
QY 903 ATGATTAACGAGAGCTCAATAGCACATATATATATATATATATATATATATATATAT 962  
DB 16005 ACCATAACTACATACACAACTTACACATATATATATATATATATATATATATATAT 16064  
QY 963 AAAAAAATTTTATTTGAATTTTTTAAACAATGTATGATATATATATATATATATATAT 1022  
DB 16065 CATACAACTTACACAT 16124  
QY 1023 ACCAATTTAACCTTTAAAAAGCATCTTAATAGGAACGAGAGTGAAGATTTCAACGA 1082  
DB 16125 ATACAT 16184  
QY 1083 AGTGTT 1088  
DB 16185 TATGTT 16190

RESULT 4  
US-08-998-416-1137/c  
Sequence 1137, Application US/08998416  
Patent No. 6239264  
GENERAL INFORMATION:  
APPLICANT: Philippsen, Peter  
APPLICANT: Pohlmann, Rainer  
APPLICANT: Steiner, Sabine  
APPLICANT: Mohr, Christine  
APPLICANT: Wendland, Jurgin  
APPLICANT: Knechtle, Philipp

APPLICANT: Rebischung, Corinne  
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPHII  
AND USES THEREOF  
NUMBER OF SEQUENCES: 1152  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6239264artis Corporation  
STREET: 3054 Cornwallis Road  
CITY: Research Triangle Park  
STATE: No. 6239264th Carolina  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/998,416  
FILING DATE: 24-DEC-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: CH 0016/97  
FILING DATE: 31-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8587  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 1137:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 636 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: PRG1692RP  
US-08-998-416-1137

Query Match 3.2%; Score 46.8; DB 4; Length 636;  
Best Local Similarity 44.6%; Pred. No. 0.017;  
Matches 274; Conservative 0; Mismatches 332; Indels 8; Gaps 2;

QY 446 TTCAAAAGAAATTAATTAATTTTTTTTCATGAACGCAATTCACCGTTCGAGAAATAT 505  
DB 636 TTTTATAAGTATTTTAACTACATCTTTTATAATATTTTATTAAATTAATAATATGAT 577  
QY 506 GCTGCTATAAATAAGTAGTCTAGTGCAGAAACAAAAATTAATATCATATAAAAGAGT 565  
DB 576 AATAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 517  
QY 566 TGTAAATTAACACCAATGTTTCGTACTACAACTTAATTTGTAATTTCTTATTCAGTCA 625  
DB 516 TATTAAATAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 457  
QY 626 CAAAAATTCCAATTTCCAATTAAGAAAAATAACGCTAGACGCGCTAAGCCCAACCTCTAAG 685  
DB 456 TTCCTATAAAGAGATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 397  
QY 686 GCTAAGTTCGAGAGGTGAAGTACGACGAAATAATATATATATATATATATATATATAT 738  
DB 396 ATAAAAATAAATAATTTTACAATATTTAAATAAATAAATAAATAAATAAATAAATAAT 337  
QY 739 ATTTTTTTTAAATAACTTTTCACATAAATTTCTTTAGGAACATATCATTTAATGTTTGA 798  
DB 336 ATTTTAATAACCAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 277  
QY 799 AAAACGTGCACATAGAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 858  
DB 276 TAAAGAAAAATAAATAATCTAATAATAATTTTAAATAAATAAATAAATAAATAAATAA 217

QY 859 TTAAGGCTTCTGATCCTCTAGTGTGAGGTGATTTTCAACGCGATGATAACGAGAAAG 918  
| | | | |  
Db 216 TAAATAGTATTCATATTAATAATATTTTATAATATATATAAATATTAATGATCAAT 157  
| | | | |  
QY 919 CTCATTAGCACATTTACTTCTAGATATTTATATTAATTAACCTTGAAAAAATATTTTATTT 978  
| | | | |  
Db 156 TAAGTAAATATATATATATATATATTAAGTATTA-ATAATCAAAATTAATTAATTA 98  
| | | | |  
QY 979 GAATTTTTTAAACAAATGATAGCATATAAATATTTTTTAAACACACACCAATTTAACCCCTTT 1038  
| | | | |  
Db 97 AATAATGATAATATAGTTTAAATATTTTAAATACCTTAAATATATTAATAAATAAAGTTTAT 38  
| | | | |  
QY 1039 AAAAAGCATCTAA 1052  
| | | | |  
Db 37 ATTAATCTTTATAA 24  
| | | | |

## RESULT 5

US-08-617-860B-32/c  
; Sequence 32, Application US/08617860B  
; Patent No. 6133506  
; GENERAL INFORMATION:  
; APPLICANT: Työfer, R., Bautor, J., Bothmann, H., Filsak, E.,  
; APPLICANT: Hwike-Grandpierre, C., Klein, B., Martini, N.,  
; APPLICANT: Miller, A., Schulte, W., Vostz, M., Walek, J.,  
; APPLICANT: Schell, J.,  
; TITLE OF INVENTION: Promoters  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Steinberg, Raskin & Davidson, P.C.  
; STREET: 1140 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/617,860B  
; FILING DATE: 01-MAR-1996  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: PCT/EP94/02950  
; FILING DATE: 05-SEP-1994  
; APPLICATION NUMBER: DE P4329951.2  
; FILING DATE: 04-SEP-1993  
; INFORMATION FOR SEQ ID NO: 32:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1850 Base pairs  
; TYPE: Nucleic acid  
; STRANDEDNESS: Double stranded  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Cuphea lanceolata  
; IMMEDIATE SOURCE:  
; LIBRARY: genomic Lambda FIX II  
; CLONE: ClEg1  
; FEATURE:  
; NAME/KEY: CAAT-Signal  
; LOCATION: 1428..1432  
; FEATURE:  
; NAME/KEY: TATA-Signal  
; LOCATION: 1553..1556  
; FEATURE:  
; NAME/KEY: Transcription start  
; LOCATION: 1585  
; FEATURE:  
; NAME/KEY: Leguminbox

; LOCATION: 1642..1657  
; FEATURE:  
; NAME/KEY: Startcodon  
; LOCATION: 1797..1799  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1797..1850  
US-08-617-860B-32  
Query Match 3.2%; Score 46.6; DB 3; Length 1850;  
Best Local Similarity 47.2%; Pred. No. 0.026;  
Matches 174; Conservative 0; Mismatches 194; Indels 1; Gaps 1;  
QY 651 AATAAAACGTAGACGGCTAAGCCACCCATCTAAGGCTTAAGTTTCGAGAGGTGAAGTAGCC 710  
| | | | |  
Db 480 AAT 421  
| | | | |  
QY 711 ACGAAAAATATGATGGTTTATTAATATGATTTTTTTAAATAACTTTCACATAAATTTCT 770  
| | | | |  
Db 420 AAAATTTTTTCAATATTTCAAAAATTCAAAAATTCACAAATATTTTAAAAACTTTTAAATATATATTTT 361  
| | | | |  
QY 771 TTAGGAACATATCATTTAATGGTTTGA AAAACGTCACATAGAAAACTAAGAACGATG 830  
| | | | |  
Db 360 AAAACCAAAAAATATTTTAAAAATTCACAAATATTTTAAAAATAAAAAACTAATTTTA 301  
| | | | |  
QY 831 AGTTGGAAAACAGAGAAACACACGCTTAAGGCTTCTTGATCCTCTAGTTGGAGGTTG 890  
| | | | |  
Db 300 AATATTTTAAAAATTTTAAAAACTAATTTTATATATTTTAAAAATTTTTTTTAAAG 241  
| | | | |  
QY 891 ATTTTCAACGCGATGATPAAACGAGAAAGCTCATTAGCACATTATTACTTAGATATTATA 950  
| | | | |  
Db 240 AATTTAAAAATATATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTCAATA 182  
| | | | |  
QY 951 ATTATAACTTTGAAAAAATATTTTGAATTTTTTAAACCAATGTATGCATATAATTTAT 1010  
| | | | |  
Db 181 ATTCAATAATATCAAAATATATCCGCAAAATTTTAAATAGCAATATTTCCCGAATA 122  
| | | | |  
QY 1011 TTTTAAAAA 1019  
| | | | |  
Db 121 TTCGTAATA 113  
| | | | |

## RESULT 6

US-08-153-563-2/c  
; Sequence 2, Application US/08153563  
; Patent No. 5693506  
; GENERAL INFORMATION:  
; APPLICANT: Rodriguez, Raymond L.  
; TITLE OF INVENTION: PROCESS FOR PROTEIN PRODUCTION  
; TITLE OF INVENTION: IN PLANTS  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: Steuart Street Tower, One Market Plaza  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: US  
; ZIP: 94105-1493  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/153,563  
; FILING DATE: 16-NOV-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Kenneth A.  
; REGISTRATION NUMBER: 31,677  
; REFERENCE/DOCKET NUMBER: 2307E-515  
; TELECOMMUNICATION INFORMATION:  
; NAME/KEY: (415) 543-9600









NAME: Gass, David A.  
REGISTRATION NUMBER: 38,153  
REFERENCE/DOCKET NUMBER: 28123/34274  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4276 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Rice (*Oryza sativa*)  
STRAIN: CV. M202  
IMMEDIATE SOURCE:  
LIBRARY: (EMBL) genomic  
CLONE: -Amy7-C  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(2459..2473, 2582..2713, 2807..3619, 3704  
LOCATION: ..3952)  
FEATURE:  
NAME/KEY: mat peptide  
LOCATION: join(2459..2473, 2582..2713, 2807..3619, 3704  
LOCATION: ..3952)  
PUBLICATION INFORMATION:  
AUTHORS: Yu et al., Su-May  
TITLE: Regulation of -amylase-encoding gene expression  
TITLE: in germinating seeds and cultured cells of rice  
JOURNAL: Gene  
VOLUME: in press  
US-09-072-435-3

Query Match 3.2%; Score 46.6; DB 4; Length 4276;  
Best Local Similarity 51.7%; Pred. No. 0.035;  
Matches 106; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 839 AACAGAGAAAAACACAGCCTTAAGCCTTCTAGTCTCTAGTGGAGGTTGATTTCAA 898  
DB 2069 AACATGAATACCTGATCGCACACAGCATTGAGAGCAGAGAAAGAAATCGATTGAGA 2010  
QY 899 ACGCATGATAAAGCAGAAAGCTCATTAGCATTATTAGATATTATTAATAATAA 958  
DB 2009 ATGACGAAACAGAGCTGAGCCATTACGTTAATTAATTAATTAATTAATTAATA 1950  
QY 959 CTTGAAAAAATATTATTATTGAAATTTTAAACAAATGATGCATAAATTTATTTTAA 1018  
DB 1949 TTTTAAAAATAGATAATAATTTTAAAGTAACCTTCCTATAGAAAATTTTGCAAAA 1890  
QY 1019 ACACACCAATTTAACCTTTAAAA 1043  
DB 1889 ATCATACCGTTTAGTAGTTTCAGGAA 1865

RESULT 12  
US-09-072-917A-3/c  
Sequence 3, Application US/09072917A  
Patent No. 6288302  
GENERAL INFORMATION:  
APPLICANT: Yu, Su-May  
APPLICANT: Liu, Li-Fei  
APPLICANT: Chan, Ming-Tsair  
TITLE OF INVENTION: Application of Alpha-Amylase Gene  
TITLE OF INVENTION: Promoter and Signal Sequence in the Production of  
Patent No. 6288302  
TITLE OF INVENTION: Recombinant Proteins in Transgenic Plants and Transgenic  
TITLE OF INVENTION: Plant Seeds  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive/6300 Sears Tower

CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/072,917A  
FILING DATE:  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/509,962  
FILING DATE: 01-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Gass, David A.  
REGISTRATION NUMBER: 38,153  
REFERENCE/DOCKET NUMBER: 28123/34257  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4276 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Rice (*Oryza sativa*)  
STRAIN: CV. M202  
IMMEDIATE SOURCE:  
LIBRARY: (EMBL) genomic  
CLONE: alpha-Amy7-C  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(2459..2473, 2582..2713, 2807..3619, 3704  
LOCATION: ..3952)  
FEATURE:  
NAME/KEY: mat peptide  
LOCATION: join(2459..2473, 2582..2713, 2807..3619, 3704  
LOCATION: ..3952)  
PUBLICATION INFORMATION:  
AUTHORS: Yu et al., Su-May  
TITLE: Regulation of alpha-amylase-encoding gene expression  
TITLE: in germinating seeds and cultured cells of rice  
JOURNAL: Gene  
VOLUME: in press  
US-09-072-917A-3

Query Match 3.2%; Score 46.6; DB 4; Length 4276;  
Best Local Similarity 51.7%; Pred. No. 0.035;  
Matches 106; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 839 AACAGAGAAAAACACAGCCTTAAGCCTTCTAGTCTCTAGTGGAGGTTGATTTCAA 898  
DB 2069 AACATGAATACCTGATCGCACACAGCATTGAGAGCAGAGAAAGAAATCGATTGAGA 2010  
QY 899 ACGCATGATAAAGCAGAAAGCTCATTAGCATTATTAGATATTATTAATAATAA 958  
DB 2009 ATGACGAAACAGAGCTGAGCCATTACGTTAATTAATTAATTAATTAATAATAA 1950  
QY 959 CTTGAAAAAATATTATTATTGAAATTTTAAACAAATGATGCATAAATTTATTTTAA 1018  
DB 1949 TTTTAAAAATAGATAATAATTTTAAAGTAACCTTCCTATAGAAAATTTTGCAAAA 1890  
QY 1019 ACACACCAATTTAACCTTTAAAA 1043  
DB 1889 ATCATACCGTTTAGTAGTTTCAGGAA 1865

	Query Match Best Local Similarity Matches 195; Conservative	3.2%; 45.9%; 0;	Score 45.8; Pred. No. 0.069; Mismatches 227;	DB 3; Length 9048; Indels 3; Gaps 1;
Qy 727	TTTATTAAATGATTTTTTTTAAATACCTTTCACATAAATTTCTTTAGGAAACATATCAT	786		
Db 1942	TATAAAAAACATAATGATTTTTTACCAACCTTTTTTCTATTCTTTTTTTTTTTTTTTTT	2001		
Qy 787	TTAATCGTTTGA AAAACGTCACATAAGAAAACCTAAGACGA-TGAGTTGGGAAACAGA	845		
Db 2002	TTTTTACTTTGAAAAA AAAAAA AAAAAA AAAAAA AAAAAA AAAAAA AAAAAA AAAAAA	2061		
Qy 846	GAAAAACACACGCTTAAGGCTCTTGATCCTCTAGTTGGAGGTTGATTTTCAAACGCCATG	905		
Db 2062	ATTACTTTGGTTTTTTTGGATTTTTTTTTTAAATAAATTTAATTTCTCTATCTAA	2121		
Qy 906	ATAAACGAGAAAGCTCATAGACATATTACTTAGATATTTAATTAATTAACCTTGAAA	965		
Db 2122	TTATACCTTATTTAATAATTTGGATAATATATCAAAATATTTATCAGTTTGGCATGACA	2181		
Qy 727	TTTATTAAATGATTTTTTTTAAATAAATTTTACACATAAATTTCTTTAGGAAACATATCAT	786		
Db 383	TATATTAAATATAACATATAAATTTTCACTGATTTAAATAAATAATTCGTTTTTAATATATAAA	442		
Qy 787	TTAATCGTTTGA AAAACGTCGACATAAGAAAACCTAAGACGA---TGAGTTGGGAAACAA	843		
Db 443	TTCAATTAAGATGATTTCTTTTAAATAACACATAGCGAACATCAGCTAGAAATTAATAATAA	502		
Qy 844	GAGAAAAACACAGCCTTAAGGCTCTTGATCCTCTAGTTGGAGGTTGATTTTCAAACGCA	903		
Db 503	ATCAATAAACCCTAGCTAAAAGTCTAAAACCTTAATAAATAATGACAAATGAAGAAAAATTAAC	562		

QY 904 TGATRAACGAGAAAGCTCATTAGCACATTATTACTAGATATTTTATTAATTAATAAAGTTGA 963  
Db 563 TAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 622  
QY 964 AAAAATAATTTTAAATTTTAAACAAATGATGATGATGATGATGATGATGATGATGATGAT 1023  
Db 623 ATAAGATTTCTTAAGGGTATTTTAAATTAATAAATAAATAAATAAATAAATAAATAAATAA 682  
QY 1024 CCAATTTAAACCCCTTTAAAGAGCATCTTAATAGGAAACGAGAGGTTAAAGATTTCACCGAA 1083  
Db 683 ATAATCAATGATTTAGCTAAACCTAATAAACAATGACAAATTTACTAAA 742  
QY 1084 GTGTTGGATATGAATAATGGGTGGGATTTAGAAATTTGGTAAATGATCAGGTTAGGAT 1143  
Db 743 ATTATTGATAATATAAATAATGATTGATTTCTTAATAACAAAATTAATAAATAAAGATTTT 802  
QY 1144 TAAAT 1148  
Db 803 TAAAT 807

RESULT 15  
US-08-998-416-1137  
; Sequence 1137, Application US/08998416  
; Patent No. 6239264  
; GENERAL INFORMATION:  
; APPLICANT: Philippsen, Peter  
; APPLICANT: Pohlmann, Rainer  
; APPLICANT: Steiner, Sabine  
; APPLICANT: Mohr, Christine  
; APPLICANT: Wendland, Jurgen  
; APPLICANT: Knechtle, Philipp  
; APPLICANT: Reibschung, Corinne  
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPHII  
; TITLE OF INVENTION: AND USES THEREOF  
; NUMBER OF SEQUENCES: 1152  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 6239264artis Corporation  
; STREET: 3054 Cornwallis Road  
; CITY: Research Triangle Park  
; STATE: No. 6239264th Carolina  
; COUNTRY: USA  
; ZIP: 27709  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/998,416  
; FILING DATE: 24-DEC-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: CH 0016/97  
; FILING DATE: 31-DEC-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meigs, J. Timothy  
; REGISTRATION NUMBER: 38,241  
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 919-541-8587  
; TELEFAX: 919-541-8689  
; INFORMATION FOR SEQ ID NO: 1137:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 636 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ORIGINAL SOURCE:  
; ORGANISM: PAG1692RP  
US-08-998-416-1137

Query Match 3.1%; Score 45; DB 4; Length 636;  
Best Local Similarity 47.8%; Pred. No. 0.045;  
Matches 164; Conservative 0; Mismatches 175; Indels 4; Gaps 1;  
QY 701 TGAAGTACGACGAGAAATATATGATGTTTATTAATATATGATTTTATTAATTAACCTTTCAC 760  
Db 274 TTAATAAATTTATTAATAATAGATTATCAATAAATTAATATATATATTTATTAATTTGTTTAA 333  
QY 761 ATAAATTTCTTTAGGAAACATATCATTTAAATGTTTGAACAAACGTGCACATAAGAAAACT 820  
Db 334 AATAATATATTTTATTTATTAAGAGATTAAATTTATTTAAATATTTGTAATTTATTTT 393  
QY 821 AGAACGATGAGTTGGGAAACAGAGAGAAACACACAGCCCTTAAGGCTTCTTGTATCTCTAG 880  
Db 394 TATTATAATATCTATTTTATATA---AATATTATGTTGATTTATATTTTAAATCTTTT 449  
QY 881 TTGGAGGTTGATTTTCAACCGCATGATAAACGAGAAAGCTCATTAGCACATTTACTTTA 940  
Db 450 ATAAGAAATTTATTAATAAATTAATTTTAACTTTTAACTTTTATTTATTTATTTATTA 509  
QY 941 GATATTTATTAATATTAACCTTGAACAAATATTTATTTGAATTTTAAACAATGTATGC 1000  
Db 510 TTTAATAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 569  
QY 1001 ATAAATTTATTTTAAACAAACACACCAATTTAACCCCTTTAAAAA 1043  
Db 570 ATATTTTATCATTTATTTAATTAATTAATAAATAATATTAATAAGA 612

Search completed: March 15, 2003, 17:00:12  
Job time: 114.584 secs

GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: March 15, 2003, 15:53:40 ; Search time 97.3949 Seconds  
(without alignments)  
10377.192 Million cell updates/sec

Title: US-09-802-937-2  
Perfect score: 1440  
Sequence: 1 cgcgaggaagattaattagg.....cgcgacgcgagtgctgcgac 1440

Scoring table: IDENTITY NUC  
Gapop 10.0 , Capext 1.0

Searched: 501302 seqs, 350932545 residues  
Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1384.6	96.2	1443	10	US-09-802-927-2
2	1384.6	96.2	4354	10	US-09-802-927-1
3	62.4	4.3	2000	10	US-09-887-576-828
4	55.8	3.9	8416	8	US-08-910-386A-4
5	55.8	3.9	13341	8	US-08-910-386A-1
6	55.8	3.9	19639	8	US-08-910-386A-6
7	53.2	3.7	5940	8	US-08-910-386A-10
8	50.4	3.5	2000	10	US-09-887-576-843
9	50.4	3.5	335913	9	US-09-754-853A-2
10	50.4	3.5	335913	9	US-09-754-853A-3
11	50.2	3.5	33675	9	US-09-921-992-2
12	48.8	3.4	335913	9	US-09-754-853A-2
13	48.8	3.4	335913	9	US-09-754-853A-3
14	47.2	3.3	840681	10	US-09-790-988-1
15	46.6	3.2	419	10	US-09-960-352-11234
16	45.4	3.2	2000	9	US-09-938-842A-4747
17	44.6	3.1	442	10	US-09-960-352-12911
18	44.6	3.1	2267	10	US-09-822-830A-227
19	44.6	3.1	19639	8	US-08-910-386A-6

20	44.4	3.1	2000	9	US-09-938-842A-3659	Sequence 3659, Ap
21	44.4	3.1	7204	8	US-08-910-386A-11	Sequence 11, Appl
22	44.2	3.1	516	10	US-09-960-352-5785	Sequence 5785, Ap
23	44.2	3.1	2000	9	US-09-938-842A-4774	Sequence 4774, Ap
24	44	3.1	2000	9	US-09-938-842A-3672	Sequence 3672, Ap
25	43.8	3.0	393	10	US-09-960-352-4582	Sequence 4582, Ap
26	43.8	3.0	3950	10	US-09-887-586A-33	Sequence 33, Appl
27	43.8	3.0	3950	10	US-09-903-012-33	Sequence 33, Appl
28	43.6	3.0	449	10	US-09-960-352-2412	Sequence 2412, Appl
29	43.2	3.0	277	10	US-09-960-352-12673	Sequence 12673, A
30	43.2	3.0	342	10	US-09-960-352-2399	Sequence 2399, Ap
31	42.8	3.0	1943	10	US-09-760-731-2	Sequence 2, Appli
32	42.8	3.0	3759	12	US-10-051-952-7	Sequence 7, Appli
33	42.8	3.0	6265	10	US-09-129-112-3	Sequence 3, Appli
34	42.8	3.0	513509	9	US-09-754-853A-4	Sequence 4, Appli
35	42.6	3.0	1152	9	US-10-098-841-163	Sequence 163, App
36	42.6	3.0	28001	9	US-10-193-295-3	Sequence 3, Appli
37	42.4	2.9	2000	9	US-09-938-842A-2830	Sequence 2830, Ap
38	42.4	2.9	3056	9	US-10-037-598-30	Sequence 30, Appl
39	42.4	2.9	3142	9	US-10-037-598-37	Sequence 37, Appl
40	42.4	2.9	3410	9	US-10-196-063-1	Sequence 1, Appli
41	42.4	2.9	74586	10	US-09-781-558-3	Sequence 3, Appli
42	42.2	2.9	684973	10	US-09-263-959-1	Sequence 1, Appli
43	42	2.9	513509	9	US-09-754-853A-4	Sequence 4, Appli
44	41.8	2.9	431	10	US-09-960-352-5558	Sequence 5558, Ap
45	41.6	2.9	446	10	US-09-960-352-3400	Sequence 3400, Ap

ALIGNMENTS

RESULT 1  
US-09-802-927-2  
; Sequence 2, Application US/09802927  
; Patent No. US20010031236A1  
; GENERAL INFORMATION:  
; APPLICANT: UCHIMIYA, HIROFUMI  
; APPLICANT: ARAI, SATOSHI  
; APPLICANT: FUSHIMI, TAKAOMI  
; APPLICANT: TAGAWA, MICHITO  
; APPLICANT: FUKUZAWA, HIROMITSU  
; TITLE OF INVENTION: DNA FRAGMENT HAVING PROMOTER FUNCTION  
; FILE REFERENCE: 204323USCIP  
; CURRENT APPLICATION NUMBER: US/09/802,927  
; CURRENT FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: PCT/JF98/04088  
; PRIOR FILING DATE: 1988-09-10  
; PRIOR APPLICATION NUMBER: PCT/JF99/04847  
; PRIOR FILING DATE: 1999-09-08  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 1443  
; TYPE: DNA  
; ORGANISM: Oryza sativa L cv.Nipponbare  
US-09-802-927-2

Query Match	96.2%	Score 1384.6;	DB 10;	Length 1443;
Best Local Similarity	99.2%	Pred. No. 3e-265;		
Matches 1434;	Conservative	0;	Mismatches 4;	Indels 7; Gaps 4;
QY	1	CTGCAGGAAGATTAATTAGTGGACACACACCAACCCCTGTGTTGGTGGACCCCTGTTGTT	60	
Db	1	CTGCAGGAAGATTAATTAGTGGACACACACCAACCCCTGTGTTGGTGGACCCCTGTTGTT	60	
QY	61	AATCACTGGGGTGTTCCTGGACATGCTTTTTCAGGAAATTAAGCAAGAAATTAAG	120	
Db	61	AATCACTGGGGTGTTCCTGGACATGCTTTTTCAGGAAATTAAGCAAGAAATTAAG	120	
QY	121	AAGAATGCTCAAGCTGACATGAGAAACGTAATCCAATGGAAGCGAATTTCAAGTCGTTTC	180	
Db	121	AAGAATGCTCAAGCTGACATGAGAAACGTAATCCAATGGAAGCGAATTTCAAGTCGTTTC	180	





```
QY 420 ACTGAGCTGCAATACGATTTCTTTTCAAAAAGAAATTAATAATTTTTCATGAA 479
Db 421 ACTGAGCTGCAATACGATTTCTTTTCAAAAAGAAATTAATAATTTTTCATGAA 480
QY 480 AGCGAATTCACCGTTGAGAAATATCTGCTATAAATAAGTAGTCTAGTGCAGAAACAA 539
Db 481 AGCGAATTCACCGTTGAGAAATATCTGCTATAAATAAGTAGTCTAGTGCAGAAACAA 540
QY 540 AATTAATATCATATAAAAAGAGGTTGTTAATTAACAACCATGTTTGTACTACAACTC 599
Db 541 AATTAATATCATATAAAAAGAGGTTGTTAATTAACAACCATGTTTGTACTACAACTC 600
QY 600 TAAATTTGTAATTTCTTATTTTCAGTCAAAAATTCCTCAATTTCCAAATTAAGAAAAATAAACG 659
Db 601 TAAATTTGTAATTTCTTATTTTCAGTCAAAAATTCCTCAATTTCCAAATTAAGAAAAATAAACG 660
QY 660 TAGACGGCTAAGCCACCCATCTAAGCTAAGTTCGAGAGGTGAAGTAGCGACGAAAT 719
Db 661 TAGACGGCTAAGCCACCCATCTAAGCTAAGTTCGAGAGGTGAAGTAGCGACGAAAT 720
QY 720 ATGATGGTTTATTAATATGATTTTTTTTAAATAAATTTTACATAAATTTCTTTAGGAAAC 779
Db 721 ATGATGGTTTATTAATATGATTTTTTTTAAATAAATTTTACATAAATTTCTTTAGGAAAC 780
QY 780 ATATCATTTAATGTTTCAAAAAGCTGCACATAAGAACTAAGAACGATGATTTGGGAA 839
Db 781 ATATCATTTAATGTTTCAAAAAGCTGCACATA - GAAACTAGAACGATGATTTGGGAA 838
QY 840 ACAAGAGAAAAACACAGCTTTAAGGCTTCTTGATCTCTAGTTGGAGGTTGATTTTCAA 899
Db 839 ACAAGAGAAAAACACAGCTTTAAGGCTTCTTGATCTCTAGTTGGAGGTTGATTTTCAA 898
QY 900 CGCATGATAAAGCGAAGCTCATACGATTTATTAATTAATTAATTAATTAATTAATTA 959
Db 899 CGCATGATAAAGCGAAGCTCATACGATTTATTAATTAATTAATTAATTAATTAATTA 958
QY 960 TTGAAAAAATATTTATTTGAAATTTTTTAAACAATGTATGCAFAAATTTTAAAAA 1019
Db 959 TTGAAAAAATATTTATTTGAAATTTTTTAAACAATGTATGCAFAAATTTTAAAAA 1018
QY 1020 CACACCAATTTAACCTTTAAAAAGCATCTAATAGAAACGAGGAAGTTAAAGATTCCAC 1079
Db 1019 CACACCAATTTAACCTTTAAAAAGCATCTAATAGAAACGAGGAAGTTAAAGATTCCAC 1078
QY 1080 CGAA---GTGTTTGATTAATGAAATATGGGTGGGATTTAGAAATTTGTAATGAATCAGGG 1136
Db 1079 CGAAGGTGTTTGGATTAATGAAATATGGGTGGGATTTAGAAATTTGTAATGAATCAGGG 1138
QY 1137 TTAGGATTAATATTTAAATGAAAGAGGAGAAATGAATGGTTAGAGTTTAAATGTGCTTT 1196
Db 1139 TTAGGATTAATATTTAAATGAAAGAGGAGAAATGAATGGTTAGAGTTTAAATGTGCTTT 1198
QY 1197 TTTGTTGGGTAGAAAATTTATTTGGCATACACTCCCGAGAGCGGTGCGTGTGGTGGG 1256
Db 1199 TTTGTTGGGTAGAAAATTTATTTGGCATACACTCCCGAGAGCGGTGCGTGTGGTGGG 1258
QY 1257 CAGAAAGCGTCTTTTTCGTTGAAAAAATAAATCTGCTTAAAAAGGAAACAGAAAGAGCCAG 1316
Db 1259 CAGAAAGCGTCTTTTTCGTTGAAAAAATAAATCTGCTTAAAAAGGAAACAGAAAGAGCCAG 1318
QY 1317 CTTT-GTTGTACCGTCTCACCAGAACGAAACAAAAGCCCAACCACTTAACTCTC 1375
Db 1319 CTTTGGTTGTACCGTCTCACCAGAACGAAACAAAAGCCCAACCACTTAACTCTC 1378
QY 1376 GATCGACCGAGACTCTCTCCATTTTCAGGGCGCACGCGAGAGACGCGACGCGAGTTTCG 1435
Db 1379 GATCGACCGAGACTCTCTCCATTTTCAGGGCGCACGCGAGAGACGCGACGCGAGTTTCG 1438
QY 1436 TCGAC 1440
Db 1439 TCGAC 1443
```

```
RESULT 3
US-09-887-576-828/c
; Sequence 828, Application US/09887576
; Patent No. US2002014047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001US1
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 828
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-887-576-828

Query Match 4.3%; Score 62.4; DB 10; Length 2000;
Best Local Similarity 61.9%; Pred No. 0.0017;
Matches 99; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 903 ATGATAACGAGAAAGCTCATTAGCACATTTACTTAGATATTTATTAATTAACCTTG 962
Db 1227 ACGTAAACGAGATGACCCATTAACGTATGATTAATTAAGTATTAACCTTTTAAACTTT 1168
QY 963 AAAAAATATTTATTTGAATTTTTTAAACAATGTATGCATAAAATTTATTTTAAAAACAC 1022
Db 1167 AAAAAATAGATTATATGATTTTTTAAAGCAACTTTCTCTATAGAAATTTTTCGCAAAAAAC 1108
QY 1023 ACCAATTTAACCTTTAAAGCATCTTAATAGGAACGA 1062
Db 1107 ACACCGTTTAGTTCGAAACGTCGCGTGGAATGA 1068

RESULT 4
US-08-910-386A-4/c
; Sequence 4, Application US/08910386A
; Patent No. US20020092041A1
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Hulbert, Scot
; APPLICANT: Richter, Todd
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
;
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910.386A
; FILING DATE: 13-AUG-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058950US
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8416 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORGANISM: Oryza longistaminata
; STRAIN: IRBB21
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 11
; MAP POSITION: 11q, RG103
; FEATURE:
; NAME/KEY: CDS
; LOCATION: Join(4771..7384, 7676..8052)
; OTHER INFORMATION: /product= "receptor kinase-like protein"
; OTHER INFORMATION: /note= "Xa21 gene family member A1"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 7432..7614
; OTHER INFORMATION: /note= "Snap-O11, transposon-like"
; OTHER INFORMATION: element"
;
; US-08-910-386A-4
;
; Query Match 3.9%; Score 55.8; DB 8; Length 8416;
; Best Local Similarity 57.0%; Pred. No. 0.05;
; Matches 102; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
;
; QY 891 ATTTCACGCGATGATAACGAGAAAGCTCATTAGCACCATTATTACCTAGCATATTATATATATA 950
; DB 4246 AATTGACACGATAGCTAGACGAAATAGCTGTAGTGTAATTAATTAGTAGTATTT 4187
; QY 951 ATTATAACTTGAAAAATATTTATTTGAATTTTAAACAAATGATATGCATAAATTATT 1010
; DB 4186 ATTTAACTTAAAAAATAGATTAATATTTTAAAGAAACTTTTCTATATAAAGTTT 4127
; QY 1011 TTTTAAAAACACACCAATTTAACCTTTTAAAAAGCATCTAATAGCAACGAGGAAGTT 1069
; DB 4126 TTTAAAAAACACACCATTTACAGTTCTGAAAACATACGCGTGAACCAACGATGCTTTT 4068
;
; RESULT 5
; US-08-910-386A-1/c
; Sequence 1, Application US/08910386A
; Patent No. US20020092041A1
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Hulbert, Scot
; APPLICANT: Richter, Todd
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910.386A
; FILING DATE: 13-AUG-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058950US
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13341 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORGANISM: Oryza longistaminata
; STRAIN: IRBB21
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 11
; MAP POSITION: 11q, RG103
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2367..4205
; OTHER INFORMATION: /product= "receptor kinase-like protein"
; OTHER INFORMATION: /note= "Xa21 gene family member D"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 4201..9071
; OTHER INFORMATION: /note= "retrofit, a copia-like,
; OTHER INFORMATION: transposon-like element"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4484..8821
; OTHER INFORMATION: /product= "retrofit"
; OTHER INFORMATION: /gene= "gag/pol"
; FEATURE:
; NAME/KEY: intron
; LOCATION: 9915..11712
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 10020..10975
; OTHER INFORMATION: /note= "Krispie, transposon-like
; OTHER INFORMATION: element"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 12626..12750
; OTHER INFORMATION: /note= "Pop-O12, transposon-like
; OTHER INFORMATION: element"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 13040..13248
; OTHER INFORMATION: /note= "Ds-rice2, transposon-like
; OTHER INFORMATION: element"
;
; US-08-910-386A-1
;
; Query Match 3.9%; Score 55.8; DB 8; Length 13341;
; Best Local Similarity 57.0%; Pred. No. 0.05;
; Matches 102; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
;
; QY 891 ATTTCACGCGATGATAACGAGAAAGCTCATTAGCACCATTATTACCTAGCATATTATATA 950
; DB 1842 AATTGACACGATAGCTAGACGAAATAGCTGTAGTGTAATTAATTAGTAGTATTT 1783
; QY 951 ATTATAACTTGAAAAAATATTTATTTGAATTTTAAACAAATGATATGCATAAATTATT 1010
```

```
Db 1782 ATTTAACTTAAATAAGATTAATATATTTTAAAGAACTTTTCTATAGAAAGTTT 1723
QY 1011 TTTTAAAAACACACCACTTTAAACCTTTTAAAGCACTCTTAATAGGAAACGAGGAAGTT 1069
Db 1722 TTTCAAAAACACACCACTTTAAACAGTTCTGAAACATACCGGTGAAACAGATGCTTTT 1664

RESULT 6
US-08-910-386A-6/c
; Sequence 6, Application US/08910386A
; Patent No. US20020092041A1
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Hulbert, Scot
; APPLICANT: Richter, Todd
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,386A
; FILING DATE: 13-AUG-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058950US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0300
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19639 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Oryza longistaminata
; STRAIN: IRBB21
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 11
; MAP POSITION: 11q, RG103
; FEATURE:
; NAME/KEY: -
; LOCATION: 5213..18201
; OTHER INFORMATION: /note= "Xa21 gene"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(5213..7889, 8732..9132)
; OTHER INFORMATION: /product= "receptor kinase-like protein"
; OTHER INFORMATION: /note= "Xa21 disease resistance gene"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 9645..9769
; OTHER INFORMATION: /note= "Pop-Oil1, transposon-like"
; OTHER INFORMATION: element"
; FEATURE:
; NAME/KEY: misc_feature
```

```
; LOCATION: 13040..13248
; OTHER INFORMATION: /note= "Ds-rice1, transposon-like"
; OTHER INFORMATION: element"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(15118..17720, 17827..18204)
; OTHER INFORMATION: /product= "receptor kinase-like protein"
; OTHER INFORMATION: /note= "Xa21 gene family member C; 2 bp of
; OTHER INFORMATION: deletion causing frame-shift mutation of
; OTHER INFORMATION: ORF compared to family member A1"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 16183..16184
; OTHER INFORMATION: /note= "location of 2 bp deletion
; OTHER INFORMATION: compared to family member A1"
; US-08-910-386A-6
; Query Match 3.9%; Score 55.8; DB 8; Length 19639;
; Best Local Similarity 57.0%; Pred. No. 0.061;
; Matches 102; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
QY 891 ATTTTCAAAACGATGATAAACGAGAAAGCTCATTAGCACATTATTACTTAGATATTTATA 950
Db 4698 AATTGACACAGATACGTAGAACGAAATAAGCTGTAGTGATAATTTAATTAAGTAGTAATT 4629
QY 951 ATTATTAACCTTGAANAATATTTATTGAAATTTTAAACAGTGTATGCATAAATTATT 1010
Db 4628 ATTTTAACTTTAAANAATAGATTAATATATTTTAAAGAAACITTTCTTAGAAAGTTT 4569
QY 1011 TTTTAAAAACACACCAATTTTAACCTTTTAAAGCATCTTAATAGGAACGAGGAAGTT 1069
Db 4568 TTTCAAAAACACACACCATTTTAACAGTTCTGAAACATACCGGTGAAACCGATGCTTTT 4510

RESULT 7
US-08-910-386A-10/c
; Sequence 10, Application US/08910386A
; Patent No. US20020092041A1
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Hulbert, Scot
; APPLICANT: Richter, Todd
; TITLE OF INVENTION: Procedures and Materials for Conferring
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,386A
; FILING DATE: 13-AUG-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058950US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0300
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5940 base pairs
```





```
; Patent No. US20020069426A1
; GENERAL INFORMATION:
; APPLICANT: Boronat, Albert;
; APPLICANT: Campos, Narciso;
; APPLICANT: Rodriguez-Concepcion, Manuel;
; APPLICANT: Rohmer, Michel;
; APPLICANT: Seeman, Myriam;
; APPLICANT: Valentin, Henry E.;
; APPLICANT: Venkatesh, Tyamagondlu V.;
; APPLICANT: Venkatesh, Mylavazapu
; TITLE OF INVENTION: Methyl-D-Erythritol Phosphate Pathway Genes
; FILE OF INVENTION: 16516.107/35-21(51897)US
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 09/921,992
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 85
; SEQ ID NO 2
; LENGTH: 33675
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6924)..(7019), (7163)..(7269), (7344)..(7444), (7525)..(7634),
; LOCATION: (7694)..(7813), (7923)..(8153), (8253)..(8369), (8515)..(8589),
; LOCATION: (9012)..(9071), (9163)..(9225), (9328)..(9472), (9589)..(9750),
; LOCATION: (9911)..(10028), (10134)..(10293), (10694)..(10798),
; LOCATION: (11028)..(11129)
; NAME/KEY: unsure
; LOCATION: (1..33675)
; OTHER INFORMATION: unsure at all n locations
US-09-921-992-2

Query Match 3.5%; Score 50.2; DB 10; Length 33675;
Best Local Similarity 66.4%; Pred. No. 0.9;
Matches 87; Conservative 0; Mismatches 43; Indels 1; Gaps 1;

Qy 922 ATTGACATATTACTAGATATTTATATTAATTAACCTTGAAAAATATTTATTTGAA 981
Db 25132 AATAACTTTTGATTAATAGATTTAATTAATTAACAACTTAAAAAAGAGTAATCTGAT 25191

Qy 982 TTTTAAACAAGTATGCAATTAATTTTAAACAACACAC-CAATTTAAACCTTTAA 1040
Db 25192 ATTTTAAACAACCTTCATATAGAAATTTTACAGAAACGCCCGTTTAAACAGTTGA 25251

Qy 1041 AAAGCATCCTA 1051
Db 25252 AAAGCGTGCCA 25262

RESULT 12
US-09-754-853A-2/c
; Sequence 2, Application US/09754853A
; Publication No. US20030005491A1
; GENERAL INFORMATION:
; APPLICANT: Hauge, Brian M.
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Parsons, Jeremy D.
; APPLICANT: Wang, Ming Li
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated with
; FILE OF INVENTION: Soybean Cyst Nematode Resistance
; FILE REFERENCE: 38-10(15810)B
; CURRENT APPLICATION NUMBER: US/09/754,853A
; CURRENT FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/174,880
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 1119
; SEQ ID NO 2
; LENGTH: 335913
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: CDS

; LOCATION: (45163)..(45314), (45450)..(45509), (46941)..(48763), (48975)..(49573)
; OTHER INFORMATION: Clone ID: 240017_region_G3
US-09-754-853A-2

Query Match 3.4%; Score 48.8; DB 9; Length 335913;
Best Local Similarity 45.3%; Pred. No. 3;
Matches 354; Conservative 0; Mismatches 407; Indels 21; Gaps 4;

Qy 449 AAAAAGAAATTAATTAATTTTTCATGAACGCAATTCACCGTTTCGAGAAATATGCT 508
Db 183354 AACAAAAATTAATTTCAATTTAACTAGTAATTTAAATTTGAAATTAATGCAATTAATTTT 183295

Qy 509 GTCATAAATAGTAGTCTAGTCGAGAAACAAAATTAATATCATATAAAGAAAGAGTTGT 568
Db 183294 TTAATAATTAATTTTATCACTCATATAATATTAATTTGAATTCGAATATCACGCTCATAGT 183235

Qy 569 TAATTACAAAC-----CATGTTTCGTACTCAACTCTAATTTGTAATTTCTTATT 618
Db 183234 AAAATGTAATTCCTTCCATTCCTTTTATAGAATAAATCACAAATTTTATTTCTCTTTT 183175

Qy 619 TCAGTCACAAAATTTCCAAATTTCCAAATTAAGAAAAATAAACGCTAGACGGCTAACGCCACCC 678
Db 183174 TAAGTAAATTTTGCTATTATTATTCTTTTATTAATGAGATTATATTAAAAATATCTTTCAC 183115

Qy 679 ATCTAAGGCTAAGTTCGAGAGGTGAAGTACGCGCAAGAAATATGATGTTTAAATATG 738
Db 183114 TCAACAGGTTGTTGCTTTAATATCGAGTTTAAATGCAATTAATTTAG---AAAAATATG 183058

Qy 739 ATTTTAAATAACTTTTCACTAAATTTCTTTTGAAGAAACATATCATTTAATGTTTGA 798
Db 183057 TTTATTTTAAATTAATAATTAATTTTAAATAATTAATTTTACTTAATAGTAT 182998

Qy 799 AAAA-----CGTCACATAAGAAACCTAAGAACGATGAGTTGGGAAACAAGAGAAAC 852
Db 182997 GAAATTAATTTCTTTCTTATAGAGAGAGAGAACTACTTGTGATTTTAAATA 182938

Qy 853 ACAGCCTTAAGGCTTCTTGATCCCTAGTTGGAGTTGATTTTCAACCGCATGATAACG 912
Db 182937 TATTTATATTCACCCCAATGGACCTTTTAAATTTTATGCTAATCCTAAAAATAAAT 182878

Qy 913 AGAAGCTCATTAACCAATTAATTAATTTAGATATTTAATAATTAATAAATTTGAAAAAATAT 972
Db 182877 TGAACAATAATTTTATTTATTTATTAAGAAATTTAAATGTTATGATCTAAAAAGAAGTC 182818

Qy 973 TTAATGAAATTTTAAACAATGATGCAATTAATTTTAAACAACACACCAATTTAA 1032
Db 182817 TGGTGTGAATCTATTAAAAAAGTAAGAAATTAAGTAATTTAATAATACCAATGTAC 182758

Qy 1033 CCCTTTAAAAAGCATCCTTAATAGGAAACGAGGAAGTTAAAGATTTACCGAAGTGTGGA 1092
Db 182757 AAAGCACTAAGCAAGAAACAAAGAGGATTAAGGAAAGGATGTACAAAAATAGAAAAAGA 182698

Qy 1093 TAATGAAAAATGGGTGGATTAAGTAATTAAGTAATTAAGTAATTAAGTAATTAATTA 1152
Db 182697 TTACAGGAATGAAGAGAAAGAAAT--GTAGAGAGCAATGTACACATTAAGAGNA 182640

Qy 1153 AAATGAAGAGGAGAGTAATGATGTTAGATTTTAAATGTTCTTTTGGTGGGTGAGAAA 1212
Db 182639 AAAATATAAGAGAAAAAATGTATGTATACAGAAATAGGTATTAATTTGGTAAATGGA 182580

Qy 1213 TT 1214
Db 182579 TT 182578

RESULT 13
US-09-754-853A-3/c
; Sequence 3, Application US/09754853A
; Publication No. US20030005491A1
; GENERAL INFORMATION:
; APPLICANT: Hauge, Brian M.
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Parsons, Jeremy D.
; APPLICANT: Wang, Ming Li
```

APPLICANT: Wang, Ming Li  
TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With  
FILE OF INVENTION: Soybean Cyst Nematode Resistance  
FILE REFERENCE: 38-10(15810)B  
CURRENT APPLICATION NUMBER: US/09/754,853A  
CURRENT FILING DATE: 2001-01-05  
PRIOR APPLICATION NUMBER: US 60/174,880  
PRIOR FILING DATE: 2000-01-07  
NUMBER OF SEQ ID NOS: 1119  
SEQ ID NO 3  
LENGTH: 335913  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (46798)..(48763).(48975)..(49573)  
OTHER INFORMATION: Clone ID: 240017\_region\_G3  
US-09-754-853A-3

Query Match 3.4%; Score 48.8; DB 9; Length 335913;  
Best Local Similarity 45.3%; Pred. No. 3;  
Matches 354; Conservative 0; Mismatches 407; Indels 21; Gaps 4;

QY 449 AAAAAGAAATTAATTAATTTTTTTTCATGAAACGCAATTCACCGTTCGAGAAATATGCT 508  
Db 183354 AACAAAAAATTTTCAATTTAACTAGTAATTTAAATTTGAATTCATTAATATTTT 183295

QY 509 GTCATAAATAAGTAGCTAGTCGAGAAACAAATTAATATCACAATAAAAGAGGTGT 568  
Db 183294 TTAATAAATAATTTTACTCTCATATAATAATTAATTTGAATTTGAATTCACGGCTGATG 183235

QY 569 TAATTACAAAC-----CATGTTTCGTACTACAACTCAATTTGTAAATCTTTAT 618  
Db 183234 AAAATGATTTCTTCATCTCTCTTTATAAGAAATAAATCACAATTTTATTTCTCTTTT 183175

QY 619 TCAGTCACAAAATTCCAATTTCCAAATTAAGAAAATAAACGTAGACGGCTAAGCCACCC 678  
Db 183174 TAAGTAAATTTTGTCTATTTTACTTTTAAATGAGATTAATTTAAATATCTTTTCAC 183115

QY 679 ATCTAAGGCTAAGTTTCGAGAGTGAAGTACGCACGAAAAATATGATGTTTATTAATATG 738  
Db 183114 TCAACAGGTGTTTCTTAATATCGAGTTTAAATGCAATAAATTAG---AAAAATATG 183058

QY 739 ATTTTTTTAAATACTTTCACATAAATTTCTTTAGGAAACATATCAATTAATGGTTTGA 798  
Db 183057 TTTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 182998

QY 799 AAAAA-----CGTCACATAAGAAACTAAGAACGATGAGTTGGAAACAGAGAAAC 852  
Db 182997 GAAATTAATCTTTTCTTTATAGAGAGAGAGGAGTACTTGTGATTTATTAATAATA 182938

QY 853 ACAGCCTTAAGGCTTCTGATCCTCTAGTTGGAGGTGAATTTCAAACGCAATGATAACG 912  
Db 182937 TATTTATATTCACCCCAATGACCTTTTAAATTTTATGTTCTAATCTAAAAATAAAT 182878

QY 913 AGAAGCTCATAGACATTTATCTAGATATTTATAATTAATAAATCTGAAAAAATAT 972  
Db 182877 TGAACAATATTTATTTATTTATAGATTTAAATGTTATGATCTAAGAAAGAGTC 182818

QY 973 TTAATTTGAATTTTAAACAATGTATGCATAAATTAATTTTTTAAACACACCAATTTAA 1032  
Db 182817 TGGTGTGAATCTATTAAAAAAGTAAGAAAAATTAAGTAAATTAATAACACCAATGTAC 182758

QY 1033 CCCTTTAAAAAGCATCTTAATAGGAAACGAGCAAGTTAAAGATTCACCGAAGTGTGGA 1092  
Db 182757 AAGCAACTTAAGCAAGAACAAAGAGTAAGAAAGGGATGTACAAATAGAAAAAGA 182698

QY 1093 TAATGAAAAATGGGTGGGATAGAAATTTGGTAAATGAATCAGGGTTAGGATTAATAATTA 1152  
Db 182697 TTACGAGGAAATGAAAGGAAGAAAAAT--GTAGAAGAGCAATGTACAACATAAGAGAA 182640

QY 1153 AAATGAAGAGGGAGAAATGAATGGTTAGAGTTTAAATGTGCTTTTGTGGGTAGAAA 1212

Db 182639 AAAATATAAGGAAAAAAATGTATATGTATACAGAAATAGTATAATGTTGTTGAATCGAA 182580

QY 1213 TT 1214  
Db 182579 TT 182578

RESULT 14  
US-09-790-988-1  
Sequence 1, Application US/09790988  
Patent No. US20020127687A1  
GENERAL INFORMATION:  
APPLICANT: SHIGENOBU, SHUJI  
APPLICANT: WATANABE, HIDEKI  
APPLICANT: HATTORI, MASAHIRO  
APPLICANT: SAKAKI, YOSHIYUKI  
TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS  
FILE REFERENCE: 081356/0159  
CURRENT APPLICATION NUMBER: US/09/790,988  
CURRENT FILING DATE: 2001-02-23  
PRIOR APPLICATION NUMBER: JP2000-107160  
PRIOR FILING DATE: 2000-04-07  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 640681  
TYPE: DNA  
ORGANISM: Buchnera sp.  
US-09-790-988-1

Query Match 3.3%; Score 47.2; DB 10; Length 640681;  
Best Local Similarity 48.4%; Pred. No. 7.3;  
Matches 199; Conservative 0; Mismatches 208; Indels 4; Gaps 2;

QY 245 TGTGATACTGACTTGTGTCACAAAGCATCAAAATTCCTTCTGGAGTATCTTTATCCGA 304  
Db 580300 TATACTACTGATTTATTAATAAAAAACAAGACATTTAAGTTATAATAAATTTCTTTA 580359

QY 305 AAACCCCAAGATTAATTCATTCACCTCAGGTAATTTGCTGAACTATGCAATGAATA 364  
Db 580360 AAGAAAAATCTTTAGTTAATAAAAAACAATAATCAATCATTTATGTTAGTTCATA 580419

QY 365 CAATTCGCAATATCATGGTTATCTATCTGCTCAAAATTTGAATTTGAGTCCAACTGA 424  
Db 580420 AAAAAATCTAGAAATTTGAT--TTTTTTTATTAATAAAATATGAAATTTTACTTAACT 580477

QY 425 GACTGCAATACGATTTTCTTTTCAAAAAGAAATTAATAATTTTTCATGAAACGCA 484  
Db 580478 GTATATGTTAAAAAATTAATTAATAAAAAATAAATAAATAAATAAATAAATAAATA 580537

QY 485 ATTCAACCGTTCGAGAAATATCTGTCATAAA--TAAGTAGCTAGTGCAGAAACAAAT 542  
Db 580538 ATAAAAAATTAATTTGTTATAGTAAATATATCTTAAAGTTTATTTGATTTGAAAAAAT 580597

QY 543 TAATATACATAAAAAAGAGGTTGTTAATTAACAAACCATGTTTCGTACTCAACTCTAA 602  
Db 580598 AATATTAAGAAATATAAATAATTTTACATTTGAAAAACAGATCAGAAAAATATATTA 580657

QY 603 TTTGTAAATTTCTTTTCTAGTCACAAAATTTCAATTTCCATTTAAGAAAAA 653  
Db 580658 TTTTATCTAAACCTTAATTTGAAATAATTAATAAATAATTAATAAATAAATAAATAA 580708

RESULT 15  
US-09-960-352-11234/c  
Sequence 11234, Application US/09960352  
Patent No. US20020137139A1  
GENERAL INFORMATION:  
APPLICANT: Warren, Wesley C.  
APPLICANT: Tao, Ningbing  
APPLICANT: Byatt, John C.  
APPLICANT: Mathialagan, Nagappan  
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND

; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION  
; FILE REFERENCE: 16511.006/37-21(10298)C  
; CURRENT APPLICATION NUMBER: US/09/960,352  
; CURRENT FILING DATE: 2001-09-24  
; NUMBER OF SEQ ID NOS: 15112  
; SEQ ID NO 11234  
; LENGTH: 419  
; TYPE: DNA  
; ORGANISM: Bos taurus  
; OTHER INFORMATION: Clone ID: 48-LIB3058-052-Q1-K1-D8  
US-09-960-352-11234

Query Match 3.2%; Score 46.6; DB 10; Length 419;  
Best Local Similarity 48.3%; Pred. No. 1.6;  
Matches 130; Conservative 0; Mismatches 139; Indels 0; Gaps 0;  
QY 892 TTTTCARACGCATGATAAACGAGAAAGCTCATTAGCACATTATTACTTAGATATTATATAA 951  
DB 295 TTTTAAAAAATTTTAAAAATTTTAAATTTTAAATTTTAAAAAATATAATATAAAAAATTTT 236  
QY 952 TTATAAACTTCAAAAAAATATTATTGAATTTTAAACAATGTATGCATAAATATTATTT 1011  
DB 235 TTTTATATTTAATAAAATTTATAAATTTTAACTTTTAAAAATTTTATTTATTTATTTTA 176  
QY 1012 TTTAAAAACACACCAATTTAACCCCTTTAAAAAGCATCCTAATAGGAAACGAGGAAGTTAA 1071  
DB 175 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 116  
QY 1072 AGATTCACCGAAGTGTTTGGATAATGAAAAATGGGGTGGGATTAGAAATGGTAAATGAAT 1131  
DB 115 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 56  
QY 1132 CAGGCTTAGGATTAAATATTAAAAATGAAA 1160  
DB 55 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 27

Search completed: March 15, 2003, 20:58:48  
Job time : 1764.39 secs